

Asymptotic distribution of the numbers of vertices and arcs of the giant strong component in sparse random digraphs

Boris Pittel*

Department of Mathematics
The Ohio State University
Columbus, Ohio, U.S.A.

bgp@math.osu.edu

Daniel Poole*

Department of Mathematics
The Ohio State University
Columbus, Ohio, U.S.A.

poole@math.osu.edu

May 22, 2015

Abstract

Two models of a random digraph on n vertices, $D(n, \text{Prob}(\text{arc}) = p)$ and $D(n, \text{number of arcs} = m)$ are studied. In 1990, Karp for $D(n, p)$ and independently T. Łuczak for $D(n, m = cn)$ proved that for $c > 1$, with probability tending to 1, there is a unique strong component of size of order n . Karp showed, in fact, that the giant component has likely size asymptotic to $n\theta^2$, where $\theta = \theta(c)$ is the unique positive root of $1 - \theta = e^{-c\theta}$. In this paper we prove that, for both random digraphs, the joint distribution of the number of vertices and number of arcs in the giant strong component is asymptotically Gaussian with the same mean vector $n\boldsymbol{\mu}(c)$, $\boldsymbol{\mu}(c) := (\theta^2, c\theta^2)$ and two distinct 2×2 covariance matrices, $n\mathbf{B}(c)$ and $n[\mathbf{B}(c) + c(\boldsymbol{\mu}'(c))^T(\boldsymbol{\mu}'(c))]$. To this end, we introduce and analyze a randomized deletion process which determines the directed $(1, 1)$ -core, the maximal digraph with minimum in-degree and out-degree at least 1. This $(1, 1)$ -core contains all non-trivial strong components. However, we show that the likely numbers of peripheral vertices and arcs in the $(1, 1)$ -core, those outside the largest strong component, are of polylog order, thus dwarfed by anticipated fluctuations, on the scale of $n^{1/2}$, of the giant component parameters. By approximating the likely realization of the deletion algorithm with a deterministic trajectory, we obtain our main result via exponential supermartingales and Fourier-based techniques.

Keywords: random digraphs; central limit theorem; core; deletion algorithm

*The authors gratefully acknowledge support from NSF grant # DMS-1101237.

1 Introduction

1.1 Definitions and Main Results

In this paper, we will be studying the random digraphs $D(n, m)$ and $D(n, p)$. $D(n, m)$ is uniformly distributed on all digraphs with vertex set $[n] = \{1, 2, \dots, n\}$ and $m = m(n)$ arcs. The (Bernoulli) digraph $D(n, p)$ is a random digraph on $[n]$, where each of the $n(n-1)$ possible arcs is present with probability $p = p(n)$ independently of all other arcs. As customary, we say that for a given m (p resp.) some graph property holds for $D(n, m)$ ($D(n, p)$ resp.) *asymptotically almost surely*, denoted a.a.s., if the probability that $D(n, m)$ ($D(n, p)$ resp.) has this property tends to 1 as $n \rightarrow \infty$. A digraph is *strongly connected* if for any pair of vertices v and w , there is a path from v to w and a path from w to v . A *strong component* of a digraph is a maximal strongly connected subgraph.

The phase transition in $D(n, p)$ and in $D(n, m)$ was established by Karp [18] and T. Łuczak [22], respectively. For instance, if $\lim np = c < 1$, then a.a.s. the number of vertices in the largest strong component is bounded in probability, and if $\lim np = c > 1$, then a.a.s. there is a strong giant component on $(\theta(c)^2 + o(1))n$ vertices, where θ is the unique root in $(0, 1)$ of $1 - \theta = e^{-c\theta}$. In the latter case, all other strong components have size bounded in probability. Later, T. Łuczak and Seierstad [24] investigated the size of the strong components for the cases $np = 1 \pm \epsilon$, where $\epsilon \rightarrow 0$, but $n^3\epsilon \rightarrow \infty$. In particular, for $np = 1 + \epsilon$, they demonstrated that a.a.s. the size of the giant component is $(\theta(np)^2 + o(\epsilon^2))n$. Our main result is the asymptotic joint distribution of the numbers of vertices and arcs in the strong giant component, both in $D(n, m = c_n n)$ and $D(n, p = c_n/n)$. Throughout the paper c_n satisfies the condition $\lim c_n = c \in (1, \infty)$.

Theorem 1.1. (i) *Let V_1, A_1 denote the vertex set and arc set of the largest strong component of $D(n, m)$. Suppose $m = c_n n$. Denote $\theta_n = \theta(c_n)$. Then there exists a continuous 2×2 matrix $\mathbf{B} = \mathbf{B}(c)$ such that*

$$\left(\frac{|V_1| - \theta_n^2 n}{n^{1/2}}, \frac{|A_1| - c_n \theta_n^2 n}{n^{1/2}} \right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{B}),$$

where $\mathcal{N}(\mathbf{0}, \mathbf{B})$ is the 2-dimensional Gaussian distribution with mean vector $\mathbf{0}$ and covariance matrix \mathbf{B} .

(ii) *Let V_1^p, A_1^p denote the vertex set and arc set of the largest strong component of $D(n, p)$. Suppose $p = c_n/n$. Denote $\boldsymbol{\mu}(c) = (\theta^2(c), c\theta^2(c))$. Then for the 2×2 matrix $\mathcal{B}(c) = \mathbf{B}(c) + c(\boldsymbol{\mu}'(c))^T(\boldsymbol{\mu}'(c))$, we have that*

$$\left(\frac{|V_1^p| - \theta_n^2 n}{n^{1/2}}, \frac{|A_1^p| - c_n \theta_n^2 n}{n^{1/2}} \right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathcal{B}),$$

where $\mathcal{N}(\mathbf{0}, \mathcal{B})$ is the 2-dimensional Gaussian distribution with mean vector $\mathbf{0}$ and covariance matrix \mathcal{B} .

To prove Theorem 1.1, we develop and analyze a randomized deletion algorithm for $D(n, m = c_n n)$. In steps, we successively delete *semi-isolated* vertices, those with either

in-degree zero or out-degree zero. The terminal digraph delivered by the algorithm is the $(1, 1)$ -core, the maximal subgraph with minimum in-degree and out-degree at least 1. The core contains the strong giant component, and conjecturally, a.a.s., the rest of the core has size negligible relative to random fluctuations, of order $n^{1/2}$, of the core parameters. Guided by this intuition, we zero in on the asymptotic distribution of the number of vertices and number of arcs in the $(1, 1)$ -core, rather than the strong giant component itself. Assuming the limit distribution is Gaussian, we determine parameters of this Gaussian distribution via approximating the actual realization of the deletion process by a deterministic system of partial differential equations. Once these parameters are determined, we prove that the Fourier transform of the actual 2-dimensional vector of the number of vertices and arcs in the $(1, 1)$ -core of $D(n, m = c_n n)$ does indeed converge, pointwise, to the Fourier transform of $\mathcal{N}(\mathbf{0}, \mathbf{B})$, implying part (i) of Theorem 1.1. The part (ii) follows then immediately by using the approach from Pittel [30]. To finish the proof of Theorem 1.1, we justify our above conjecture showing that in fact the difference between the $(1, 1)$ -core and the strong giant component has size of polylog order (Theorem 1.2).

In the next subsections, we relate our study to the known distributional results for the random *undirected* graphs. We also provide a preliminary insight into the salient points of our argument for the directed graphs, in the hope that it will serve as a helpful guide through the detailed proofs that follow.

1.2 Similarities to, and distinctions from the undirected case

Let $V_1(G)$ denote the vertex set of the largest component of undirected graph G . Back in 1970, Stepanov [38] proved that if $p = c/n$, $c > 1$, then $|V_1(G(n, p))|$, properly centered and scaled, is asymptotically normal. Twenty years later, Pittel [30] proved, for $G(n, m = cn/2)$, $c > 1$, a functional limit theorem for the counts of tree components of all various sizes, and used this theorem to prove asymptotic normality of $|V_1|$ for $G(n, m = cn/2)$. Stepanov's result for $G(n, p = c/n)$ followed then without much effort. Subsequently, Pittel and Wormald [36] found an asymptotic formula for the count of sparse connected 2-cores with given numbers of vertices and edges. They used this formula to prove, for $G(n, m)$, the 3-dimensional *local* Gaussian limit theorem, whence the integral limit theorem, for the three leading parameters: the number of vertices and the number of edges in the 2-core of the giant component, and the total size of trees rooted at the core vertices. The counterpart of this 3-dimensional limit theorem for $G(n, p)$ followed via the method in [30]. Ding, Kim, Lubetzky and Peres [12] obtained sharp asymptotic results for the “young” giant component, i.e. for $p = (1 + \epsilon)/n$, ($\epsilon \rightarrow 0$, $\epsilon^3 n \rightarrow \infty$). Extending to hypergraphs, Karoński and T. Łuczak [17], Behrisch, Coja-Oghlan and Kang [3], [4], Bollobás and Riordan [10] proved versions of the central limit theorem for the joint distribution of the number of vertices and number of hyperedges in the largest component of the random d -uniform hypergraph models $H_d(n, p)$ and $H_d(n, m)$, for $p = c(d-1)!/n^{d-1}$ and $m = cn/d$, where $c > 1/(d-1)$. Recently, Seierstad [37] found that for a family of random graph processes, the order of the giant component is

asymptotically normal, provided certain general conditions are met. This account is by no means complete. There have been obtained fine asymptotic results regarding such parameters of the giant component as its diameter and the mixing time of the walk on the giant component close to its inception; see, for instance, [12], and further references therein.

Recently, there has been interest in re-deriving results about the size of the giant component by analyzing “exploration” search processes which determine all the components in a graph. By analyzing a depth-first search version in $G(n, p)$, Krivelevich and Sudakov [21] found a simple proof of the phase transition. In fact, they found that for $p = c/n, c > 1$, a.a.s. $G(n, p)$ contains a path of linear length. Nachmias and Peres [27] analyzed a similar exploration process to rederive Bollobás’ [7] and Łuczak’s [23] concentration results about the size of young giant component. Barraez, Boucheron and De La Vega [2] and Bollobás and Riordan [9] proved that the size of the giant component in $G(n, p = c/n), c > 1$, is asymptotically normal using exploration processes.

In light of this progress, lack of distributional results for the strong giant component in $D(n, m)$ and $D(n, p)$ seems rather striking. In fact, just counting strongly connected digraphs had been an open challenge. Even though Bender, Canfield and McKay [5] were already able in 1992 to determine the asymptotic count of connected undirected graphs with given numbers of vertices and edges, a counterpart of their remarkable formula for the strongly connected digraphs was obtained only very recently, see Pérez-Giménez and Wormald [28], Pittel [32].

Perhaps one of the reasons for this disparity is that determining the strong components of a digraph is algorithmically more difficult than finding the components of a graph. For a digraph, the component notion morphs into two, harder-to-handle, dual notions of a sink-set and a source-set, the subsets of vertices with no arc going outside, and no arcs coming from outside, respectively. A digraph is strongly connected precisely when the full vertex set is the only source-set and the only sink-set. As a consequence, finding the strong component containing some generic vertex v depends on “global” information, i.e. this determination requires information about the digraph parts possibly quite distant from v . If we try to determine the strong component containing v by either a depth-first or breadth-first search, we apparently would have to search for vertices of 2 distinct types, the vertices that can reach v , along a directed path, and the vertices that can be reached from v . However the size of the intersection of two sets in the random digraph, in either $D(n, p)$ or $D(n, m)$, is hard to access since the two search processes are *interdependent*. We clearly need to find a middle-ground search process which would provably deliver a close *approximation* to the giant component, without us having to deal with this nasty interdependence.

We consider the following deletion algorithm on a digraph D . First, we delete all isolated vertices, those with both in-degree zero and out-degree zero, from D obtaining $D(0)$. Recursively, if $t \geq 0$ is such that $D(t)$ does not have any semi-isolated vertices, then the deletion process stops, and we define $D(s) \equiv D(t), s > t$. If $D(t)$ does have semi-isolated vertices, then

- First, we delete a semi-isolated vertex, chosen uniformly at random among all semi-

isolated vertices, along with its incident arcs from $D(t)$.

- Second, we delete all newly isolated vertices, and set $D(t+1)$ equal to the remaining subdigraph of $D(t)$.

Let $\bar{\tau}$ be the first moment that $D(t)$ does not have any semi-isolated vertex; so $D(t) = D(\bar{\tau})$ for all $t \geq \bar{\tau}$. The terminal $D(\bar{\tau})$ is both the $(1, 1)$ -core of D and the $(1, 1)$ -core of all the digraphs $D(t)$. If D has non-trivial strong components, the largest strong component is contained within the $(1, 1)$ -core. More precisely, the $(1, 1)$ -core is comprised of *all* non-trivial strong components along with directed paths between these components. Our key result is that for $D(n, m = c_n n)$, $c := \lim c_n \in (1, \infty)$, a.a.s. there are not many vertices and arcs which are in the $(1, 1)$ -core but not in the largest strong component.

1.3 Switching to the core

The following theorem allows us to switch from finding the number of vertices and arcs in the strong giant component to finding those numbers of vertices and arcs in the $(1, 1)$ -core.

Theorem 1.2. *Let $V_{1,1}$, $A_{1,1}$ denote the vertex set and the arc set of the $(1, 1)$ -core of $D(n, m = c_n n)$. Then a.a.s.*

$$\begin{aligned} 0 &\leq |V_{1,1}| - |V_1| \leq 2(\ln n)^8, \\ 0 &\leq |A_{1,1}| - |A_1| \leq 4(\ln n)^{10}. \end{aligned}$$

Theorem 1.2 calls to mind an observation that most sparse digraphs, with minimum in-out degree 1 at least, provably contain a massive strong component. This was a key ingredient in derivation of asymptotic counts of strongly connected digraphs in [28] and in [32]. Cooper and Frieze [11] used a similar property for a random digraph with a *given* degree sequence. The proof of Theorem 1.2 extends to a exploration process, the *full* depth-first search, the following observation due to Karp [18]: the size of the descendant set of a generic vertex in $D(n, p)$ has the same distribution as the size of the component containing a generic vertex in $G(n, p)$. Karp's observation has been gainfully used before, see Biskup, Chayes and Smith [6]. We will prove Theorem 1.2 in the last Section 10 since the argument does not require the properties of the deletion process. Next comes

1.4 Finding the core

Next comes

Theorem 1.3. *There is a continuous 2×2 , positive-definite, matrix $\mathbf{B} = \mathbf{B}(c)$ such that*

$$\left(\frac{|V_{1,1}| - \theta_n^2 n}{n^{1/2}}, \frac{|A_{1,1}| - c_n \theta_n^2 n}{n^{1/2}} \right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{B}),$$

where $\mathcal{N}(\mathbf{0}, \mathbf{B})$ is the 2 dimensional Gaussian distribution with mean $\mathbf{0}$ and covariance matrix \mathbf{B} .

In particular, both $|V_{1,1}|$ and $|A_{1,1}|$ undergo random fluctuations of order $n^{1/2}$ around $\theta_n^2 n$ and $c_n \theta_n^2 n$ respectively, with $n^{1/2} \gg (\ln n)^{11}$, the likely bound of the error-approximations of $|V_1|$ by $|V_{1,1}|$, and of $|A_1|$ by $|A_{1,1}|$ in Theorem 1.2. Thus Theorem 1.3 combined with Theorem 1.2 imply Theorem 1.1(i).

The $(1, 1)$ -core is a natural counterpart of the k -core in *undirected* graphs, see Bollobás [8]. The k -core of a graph is the maximal subgraph with minimum degree at least k . Pittel, Spencer and Wormald [33] introduced a randomized deletion algorithm which terminates with the k -core and used it to identify the phase transition window of width $n^{1/2+\epsilon}$ around an explicit threshold value $c_k n$ of number of edges necessary for a.a.s. existence of a k -core ($k \geq 3$), as well as to establish the likely concentration of the k -core size within $n^{1/2+\epsilon}$ distance from its expected value. Later, Janson and M. Luczak [16] undertook a distributional analysis of this deletion algorithm and proved that, for sparse $G(n, m)$, the size of the k -core is asymptotically normal with standard deviation of order $n^{1/2}$. They also demonstrated that the random moment when the Erdős-Rényi graph process $\{G(n, m)\}$ develops a k -core is asymptotically normal, again with standard deviation of order $n^{1/2}$.

To prove Theorem 1.3, we analyze the likely behavior of the deletion process for finding the $(1, 1)$ -core in the directed graph $D(n, m)$. In part, our approach has certain semblance with investigation of randomized deletion processes for the k -core problem in [33], and for the Karp-Sipser greedy matching algorithm in Aronson, Frieze and Pittel [1].

By construction, the deletion process is obviously Markovian, but prohibitively hard to analyze due to the enormous number of states. Fortunately it is possible to aggregate these states into equivalence classes preserving the Markovian nature of the process. Namely, we introduce the process $\{\mathbf{s}(t)\}$,

$$\mathbf{s}(t) = (\nu(t), \nu_i(t), \nu_o(t), \mu(t)); \quad (1.1)$$

here $\nu(t)$ is the number of vertices, $\nu_i(t)$ is the number of vertices with zero in-degree, $\nu_o(t)$ is the number of vertices with zero out-degree, and $\mu(t)$ is the number of arcs of $D(t)$. Clearly our task is to determine the asymptotic distribution of $\nu(t)$ and $\mu(t)$ at

$$\bar{t} = \min\{t : \nu_i(t) = 0, \nu_o(t) = 0\},$$

the first moment t when there are no semi-isolated vertices left. In Section 2, we show that $\{\mathbf{s}(t)\}$ indeed remains Markovian. The price for lumping together various digraphs with the same foursome \mathbf{s} is that we have to determine decidedly more involved transition probabilities. For instance, it is necessary to evaluate $g(\mathbf{s})$, the number of digraphs with a generic parameter \mathbf{s} . In Section 3, following Pittel [32], we asymptotically evaluate $g(\mathbf{s})$ for a wide range of \mathbf{s} , via an argument based on McKay's asymptotic formula [25] for the number of $(0, 1)$ -matrices with given row and column sums.

In Section 4, we introduce the joint characteristic function of $\nu(\bar{t}), \mu(\bar{t})$,

$$\varphi_{\mathbf{s}}(\mathbf{u}) = E[\exp(i\mathbf{u}^T(\nu(\bar{t}), \mu(\bar{t})))] , \quad \mathbf{u} \in \mathbb{R}^2,$$

for the deletion process that starts from a generic initial state \mathbf{s} . Due to $\{\mathbf{s}(t)\}$ being Markovian, $\varphi_{\mathbf{s}}(\mathbf{u})$ satisfies an equation

$$\varphi_{\mathbf{s}}(\mathbf{u}) = E_{\mathbf{s}}[\varphi_{\mathbf{s}'}(\mathbf{u})], \quad (1.2)$$

\mathbf{s}' standing for the random *next* state. It had been shown (e.g. Pittel [31], with Laplace, rather than Fourier transform), that an equation of this kind can be used to establish asymptotic normality in the cases when the mean and the variance of the random parameter in question are (almost) linear in n , even when no representation of this parameter as a sum of weakly dependent, uniformly negligible, terms is forthcoming. Expecting that the mean and the covariance of $(\nu(\bar{\tau}), \mu(\bar{\tau}))$ are indeed linear in n , we hope to approximate $\varphi_{\mathbf{s}}(\mathbf{u})$ by a Gaussian characteristic function, $G_n(\mathbf{s}/n, \mathbf{u})$, with a mean $n\mathbf{f}(\mathbf{s}/n) = n[f_1(\mathbf{s}/n), f_2(\mathbf{s}/n)]^T$ and and covariance matrix $n\boldsymbol{\psi}(\mathbf{s}/n) = n\{\psi_{j,k}(\mathbf{s}/n)\}$, i.e. being dependent on the initial state \mathbf{s} . Explicitly we set

$$G_n(\mathbf{s}/n, \mathbf{u}) = \exp \left(i n \sum_{j=1}^2 u_j f_j(\mathbf{s}/n) - \frac{n}{2} \sum_{j,k=1}^2 u_j u_k \psi_{j,k}(\mathbf{s}/n) \right).$$

We want to show that $|\varphi_{\mathbf{s}}(\mathbf{u}) - G_n(\mathbf{s}/n, \mathbf{u})| \rightarrow 0$, at least for “good” \mathbf{s}/n , those among the likely values of $\mathbf{s}(0)$ arising from $D(n, m)$. Assuming smoothness of $\mathbf{f}(\cdot)$ and $\boldsymbol{\psi}(\cdot)$ for good values of the argument, we wish to determine $\mathbf{f}(\cdot)$ and $\boldsymbol{\psi}(\cdot)$ out of the condition that G_n *nearly* satisfies the recurrence relation, i.e. within the additive term $o(n^{-1})$. This condition leads us to the system of first order PDE for f_i and $\psi_{j,k}$. The method of characteristics reduces the PDE to a system of ODEs, whose solution is anticipated a.a.s. to be close to the random $\{\mathbf{s}(t)\}$. In Section 5, we solve this system of equations, in explicit form for \mathbf{f} , and in integral form for $\boldsymbol{\psi}$. We do so by identifying two explicit integrals, i.e. functions of the current state, that remain constant along the characteristics of the ODE.

In Section 6, we introduce a pair of exponential super-martingales constructed from those two integrals of the ODE system and use them to show that until the end of the process, $\{\mathbf{s}(t)\}$ stays close to the deterministic trajectory, provided that $\mathbf{s} = \mathbf{s}(0)$ is “slightly better” (even closer to expected value of $\mathbf{s}(0)$ in $D(n, m)$) than just good. Therefore, for those initial \mathbf{s} , $G_n(\mathbf{s}(t)/n, \mathbf{u})$ nearly satisfies (1.2) for all $t \leq \bar{\tau}$. So, conditioned on a better-than-good initial state $\mathbf{s}(0)$, $(\nu(\bar{\tau}), \mu(\bar{\tau}))$ is asymptotically Gaussian, with mean $n\mathbf{f}(\mathbf{s}(0)/n)$ and covariance matrix $n\boldsymbol{\psi}(\mathbf{s}(0)/n)$.

However, our ultimate goal is to determine the asymptotic distribution of $(\nu(\bar{\tau}), \mu(\bar{\tau}))$ for the *random* $\mathbf{s}(0)$ in $D(n, m = c_n n)$. In Section 8, we determine the asymptotic distribution of this random $\mathbf{s}(0)$, and use it together with the limiting distribution of $(\nu(\bar{\tau}), \mu(\bar{\tau}))$ conditioned on a generic $\mathbf{s}(0)$ to prove asymptotic normality of the terminal pair $(\nu(\bar{\tau}), \mu(\bar{\tau}))$ for $D(n, m)$ as the starting state of the deletion process. We will find that the random fluctuations of $\mathbf{s}(0)$ have no influence on the limiting means, but have a discernible effect on the limiting covariance matrix.

1.5 Description of the mean and covariance parameters as $c \downarrow 1$

Although the entries of $\mathbf{B}(c)$ are in integral form, we can say much more about these entries for c close to 1, which we detail in Section 9.2. The formulas are particularly simple for the pair $(|V_1|, \text{Exc}_1)$, where $\text{Exc}_1 := |A_1| - |V_1|$ is the *excess* of the largest strong component. For $D(n, m = cn)$ and $D(n, p = c/n)$, this pair is asymptotically

Gaussian with mean $n(\theta^2, (c-1)\theta^2)$ with covariance matrices $n\tilde{\mathbf{B}}$ and $n\tilde{\mathcal{B}}$, which are determined from \mathbf{B} and \mathcal{B} respectively. For $c-1 = \epsilon \downarrow 0$, we have that $\theta = 2\epsilon + O(\epsilon^2)$ and both $\tilde{\mathbf{B}}$ and $\tilde{\mathcal{B}}$ are

$$\begin{pmatrix} 40\epsilon + O(\epsilon^2) & 60\epsilon^2 + O(\epsilon^3) \\ 60\epsilon^2 + O(\epsilon^3) & \frac{272}{3}\epsilon^3 + O(\epsilon^4) \end{pmatrix}.$$

Qualitatively this is similar to the covariance matrix of (size of 2-core, excess of 2-core) in $G(n, m = cn/2)$ and $G(n, p = c/n)$, $c > 1$, see Pittel and Wormald [36]. Note though that, unlike our present set-up, the formulas in [36] were established under a weaker condition, $\epsilon = \epsilon(n) \gg n^{-1/3}$. We conjecture that the pair $(\nu(\bar{\tau}), \mu(\bar{\tau}))$ is also asymptotically Gaussian for $\epsilon \gg n^{-1/3}$.

Comment 1.4. *Throughout this paper, all unspecified limits are always with respect to $n \rightarrow \infty$.*

2 Deletion Process

By construction, the process $\{D(t)\}$ is clearly Markovian. In this section, our goal is to show that the simpler process $\{\mathbf{s}(t)\}$, defined in (1.1), is Markovian as well. The proof below uses as a template the reduction argument in Aronson, Frieze and Pittel [1] for the Karp-Sipser [19] greedy matching algorithm.

We start with a few definitions. Let $D = (V, E)$ be a digraph. A vertex w is a *descendant* of v if either $w = v$ or there is a directed path from v to w . We call w is a *direct descendant* of v if $(v, w) \in E$. Dually, we say that w is an *ancestor* of v if either $w = v$ or there is a directed path from w to v , with w being a *direct ancestor* if $(w, v) \in E$. Let $\mathcal{S} := (V, \mathcal{O}_i, \mathcal{O}_o, \mu)$ denote the foursome composed of the vertex set of V , the set of vertices of in-degree zero, the set of vertices of out-degree zero, and the *number* of arcs.

Let us show first that $\{\mathcal{S}(t)\}$ is itself a Markov chain. While at the first step we delete all isolated vertices, at every other step we begin with a digraph without isolated vertices and deliver its sub-digraph without isolated vertices.

Substep 1. Choose a semi-isolated vertex (uniformly at random) and delete this vertex along with all incident arcs, obtaining an intermediate digraph D^* with parameter $\mathcal{S}^* = (V^*, \mathcal{O}_i^*, \mathcal{O}_o^*, \mu^*)$, along with $I^* := \mathcal{O}_i^* \cap \mathcal{O}_o^*$ being the set of newly-born isolated vertices in D^* .

Substep 2. Delete the vertices in I^* from D^* obtaining a digraph, D' , without isolated vertices, with parameter $\mathcal{S}' = (V', \mathcal{O}_i', \mathcal{O}_o', \mu')$.

The probability of a specific semi-isolated vertex being chosen is thus $1/(|\mathcal{O}_i| + |\mathcal{O}_o|)$.

We say that \mathcal{S}' *can follow* from \mathcal{S} if a digraph with parameter \mathcal{S}' can be obtained from some digraph with parameter \mathcal{S} after one step of the deletion algorithm.

Proposition 2.1. *For \mathcal{S}' to be able to follow \mathcal{S} there must exist vertex sets $A \subset \mathcal{O}_i$, $B \subset \mathcal{O}_o$, and $R_i, R_o \subset V \setminus (\mathcal{O}_i \cup \mathcal{O}_o)$ such that*

$$V' = V \setminus (A \cup B), \quad \mathcal{O}_i' = (\mathcal{O}_i \setminus A) \cup R_i, \quad \mathcal{O}_o' = (\mathcal{O}_o \setminus B) \cup R_o, \quad \mu' = \mu - k,$$

where (i) at least one of A and B have cardinality 1, (ii) at least one of R_i and R_o are empty, and (iii) $k \geq \max\{|A| + |R_o|, |B| + |R_i|\}$. Furthermore if \mathcal{S}' can follow \mathcal{S} , the sets A and B are uniquely determined by $(\mathcal{S}, \mathcal{S}')$.

Proof. For certainty, suppose we delete a vertex, v , with *in*-degree zero in substep 1. Each vertex other than v has the same out-degree in D^* as it does in D , so the vertices with out-degree zero stay the same, i.e. $\mathcal{O}_o = \mathcal{O}_o^*$. However, vertices from \mathcal{O}_o , whose only direct ancestor in the original digraph D is the deleted vertex, v , now also have zero out-degree, and hence are isolated in the intermediate digraph D^* . In fact, all isolated vertices of D^* are born this way. Non-semi-isolated vertices of D join \mathcal{O}_i^* if their only direct ancestor in D is v . In the second substep, we delete the isolated vertices, I^* , from the digraph to obtain D' . In particular $V' = V^* \setminus I^*$, $\mathcal{O}_i' = \mathcal{O}_i^* \setminus I^*$, $\mathcal{O}_o' = \mathcal{O}_o^* \setminus I^*$, $\mu' = \mu^*$.

At the end, a vertex v is deleted from \mathcal{O}_i , along with vertices, B , from \mathcal{O}_o whose only direct ancestor was v , so that $\mathcal{O}_o' = \mathcal{O}_o \setminus B$. Also, the set R_i of vertices which in D have v as their only direct ancestor, now have in-degree zero; so $\mathcal{O}_i' = (\mathcal{O}_i \setminus \{v\}) \cup R_i$. The number of arcs must decrease by at least $\max\{1, |B| + |R_i|\}$; since D had no isolated vertices, v had at least one direct ancestor, and so at least one arc is deleted in Substep 1. In particular, we have $V' = V \setminus (\{v\} \cup B)$, $\mathcal{O}_i' = (\mathcal{O}_i \setminus \{v\}) \cup R_i$, $\mathcal{O}_o' = \mathcal{O}_o \setminus B$, and $\mu - \mu' \geq |B| + |R_i|$. There is a similar description if in Substep 1 we delete a vertex of zero *out*-degree. \square

To proceed, let $\mathcal{D}_{\mathcal{S}}$ denote the set of all digraphs with parameters \mathcal{S} . For $D' \in \mathcal{D}_{\mathcal{S}'}$, let $N_{\mathcal{S}, \mathcal{S}'}^{(in)}(D')$ denote the number of digraphs $D \in \mathcal{D}_{\mathcal{S}}$ such that D' can be obtained from D via one step of this deletion algorithm by initially deleting a vertex with in-degree zero. Similarly, let $N_{\mathcal{S}, \mathcal{S}'}^{(out)}(D')$ denote the number of digraphs, D , such that D' can be obtained after deleting a vertex with zero out-degree.

Lemma 2.2. *Both $N_{\mathcal{S}, \mathcal{S}'}^{(in)}(D')$ and $N_{\mathcal{S}, \mathcal{S}'}^{(out)}(D')$ depend only on \mathcal{S} and \mathcal{S}' , i.e. there exist $N_{\mathcal{S}, \mathcal{S}'}^{(in)}$ and $N_{\mathcal{S}, \mathcal{S}'}^{(out)}$ such that*

$$N_{\mathcal{S}, \mathcal{S}'}^{(in)}(D') \equiv N_{\mathcal{S}, \mathcal{S}'}^{(in)}, \quad N_{\mathcal{S}, \mathcal{S}'}^{(out)}(D') \equiv N_{\mathcal{S}, \mathcal{S}'}^{(out)}, \quad \forall D' \in \mathcal{D}_{\mathcal{S}'}.$$

Proof. Consider, for instance, $N_{\mathcal{S}, \mathcal{S}'}^{(in)}(D')$. First of all, $N_{\mathcal{S}, \mathcal{S}'}^{(in)}(D') \equiv 0$, unless \mathcal{S} and \mathcal{S}' are such that $|A| = 1, |T| = 0$ and $\mu - \mu' \geq \max\{1, |B| + |R_i|\}$. In the latter case let v be the single vertex in A . For any starting digraph D , v is necessarily the vertex deleted in Substep 1. Furthermore, in D , each vertex in $B \cup R_i$ has v as its only direct ancestor; so v must have $|B| = |\mathcal{O}_o| - |\mathcal{O}_o'|$ arcs ending at vertices of B and $|R_i| = |\mathcal{O}_i'| - |\mathcal{O}_i| + 1$ arcs ending at vertices of R_i . Other arcs incident to v must end at a vertex in $V \setminus (\mathcal{O}_i \cup R_i \cup B)$. Hence any such D can be reconstructed by distributing these additional arcs among the vertices of $V \setminus (\mathcal{O}_i \cup R_i \cup B)$. In particular,

$$N_{\mathcal{S}, \mathcal{S}'}^{(in)}(D') = \begin{cases} \binom{|V| - |\mathcal{O}_i| - |B| - |R_i|}{\mu - \mu' - |B| - |R_i|}, & \text{if } |A| = 1, |R_o| = 0, \mu - \mu' \geq \max\{1, |B| + |R_i|\}, \\ 0, & \text{otherwise,} \end{cases}$$

implying that $N_{\mathcal{S},\mathcal{S}'}(D')$ is indeed the same for all $D' \in \mathcal{D}_{\mathcal{S}'}$. Likewise, one can show that

$$N_{\mathcal{S},\mathcal{S}'}^{(out)}(D') = \begin{cases} \binom{|V|-|\mathcal{O}_o|-|A|-|R_o|}{\mu-\mu'-|A|-|R_o|}, & \text{if } |B| = 0, |R_i| = 0, \mu - \mu' \geq \max\{1, |A| + |R_o|\}, \\ 0, & \text{otherwise.} \end{cases}$$

for all $D' \in \mathcal{D}_{\mathcal{S}'}$. □

Next

Lemma 2.3. (i) *Suppose that, conditioned on $\mathcal{S}(0)$, $D(0)$ is uniform on $\mathcal{D}_{\mathcal{S}(0)}$. Then, for each $t \geq 0$, $D(t)$ conditioned on $\mathcal{S}(0), \dots, \mathcal{S}(t)$ is distributed uniformly on $\mathcal{D}_{\mathcal{S}(t)}$.*

(ii) *Consequently, $\{\mathcal{S}(t)\}$ is Markovian.*

Proof. (i) We prove this lemma by induction on $t \geq 0$. The base case is assumed to be true by hypothesis. Suppose that, for some $t \geq 0$, conditioned on $\mathcal{S}(0), \dots, \mathcal{S}(t)$, $D(t)$ is uniformly distributed on $\mathcal{D}_{\mathcal{S}(t)}$. Let $\mathcal{S} = \mathcal{S}(t) = (V, \mathcal{O}_i, \mathcal{O}_o, \mu)$ and $\mathcal{S}' = \mathcal{S}(t+1) = (V', \mathcal{O}'_i, \mathcal{O}'_o, \mu')$. If $D' \in \mathcal{D}_{\mathcal{S}'}$, then

$$P(D(t+1) = D' | \mathcal{S}(0), \dots, \mathcal{S}(t)) = \sum_{D \in \mathcal{D}_{\mathcal{S}}} P(D(t+1) = D', D(t) = D | \mathcal{S}(0), \dots, \mathcal{S}(t)),$$

and we can break up this latter probability as follows

$$\begin{aligned} P(D(t+1) = D', D(t) = D | \mathcal{S}(0), \dots, \mathcal{S}(t)) &= \\ &P(D(t+1) = D' | D(t) = D) P(D(t) = D | \mathcal{S}(0), \dots, \mathcal{S}(t)). \end{aligned}$$

Using the inductive hypothesis, i.e. $P(D(t) = D | \mathcal{S}(0), \dots, \mathcal{S}(t)) = |\mathcal{D}_{\mathcal{S}}|^{-1}$ ($\forall D \in \mathcal{D}_{\mathcal{S}}$), we have that

$$P(D(t+1) = D' | \mathcal{S}(0), \dots, \mathcal{S}(t)) = \sum_{D \in \mathcal{D}_{\mathcal{S}}} P(D(t+1) = D' | D(t) = D) \frac{1}{|\mathcal{D}_{\mathcal{S}}|}.$$

To finish the proof of the lemma, it suffices to show that

$$\sum_{D \in \mathcal{D}_{\mathcal{S}}} P(D(t+1) = D' | D(t) = D) = \frac{N_{\mathcal{S},\mathcal{S}'}^{(in)}}{|\mathcal{O}_i| + |\mathcal{O}_o|} + \frac{N_{\mathcal{S},\mathcal{S}'}^{(out)}}{|\mathcal{O}_i| + |\mathcal{O}_o|}, \quad (2.1)$$

because if so, the probability that $D(t+1) = D'$ (conditioned on $\mathcal{S}(0), \dots, \mathcal{S}(t)$) depends only upon \mathcal{S} and \mathcal{S}' and not on choice of $D' \in \mathcal{D}_{\mathcal{S}'}$.

Now let us prove (2.1). First, we break up the event $\{D(t+1) = D'\}$ into two events depending on whether we delete a vertex with in-degree zero or out-degree zero in the first substep. Let C denote the event that we delete a vertex of in-degree zero. By symmetry, it suffices to show that

$$\sum_{D \in \mathcal{D}_{\mathcal{S}}} P(\{D(t+1) = D'\} \cap C | D(t) = D) = \frac{N_{\mathcal{S},\mathcal{S}'}^{(in)}}{|\mathcal{O}_i| + |\mathcal{O}_o|}.$$

By $N_{\mathcal{S},\mathcal{S}'}^{(in)}(D') \equiv N_{\mathcal{S},\mathcal{S}'}^{(in)}$, we know that exactly $N_{\mathcal{S},\mathcal{S}'}^{(in)}$ of these summands are non-zero. Furthermore, if D is such that the probability is non-zero, then because we know (from \mathcal{S} and \mathcal{S}' , as well as C) exactly which semi-isolated vertex is deleted first, this probability is precisely the probability of choosing this vertex in the first substep, which is $(|\mathcal{O}_i| + |\mathcal{O}_o|)^{-1}$. Part (i) is proven.

(ii) We compute

$$\begin{aligned} P(\mathcal{S}(t+1) = \mathcal{S}' | \mathcal{S}(0), \dots, \mathcal{S}(t)) &= \sum_{D' \in \mathcal{D}_{\mathcal{S}'}} P(D(t+1) = D' | \mathcal{S}(0), \dots, \mathcal{S}(t)) \\ &= \sum_{D' \in \mathcal{D}_{\mathcal{S}'}} \sum_{D \in \mathcal{D}_{\mathcal{S}(t)}} P(D(t+1) = D' | D(t) = D) \cdot P(D(t) = D | \mathcal{S}(0), \dots, \mathcal{S}(t)). \end{aligned}$$

By part (i) of the lemma, conditioned on $\mathcal{S}(0), \dots, \mathcal{S}(t)$, $D(t)$ is uniformly distributed, so

$$P(\mathcal{S}(t+1) = \mathcal{S}' | \mathcal{S}(0), \dots, \mathcal{S}(t)) = \sum_{\substack{D' \in \mathcal{D}_{\mathcal{S}'} \\ D \in \mathcal{D}_{\mathcal{S}(t)}}} P(D(t+1) = D' | D(t) = D) \cdot |\mathcal{D}_{\mathcal{S}(t)}|^{-1}, \quad (2.2)$$

which, by (2.1), depends on $\mathcal{S}(t)$ and \mathcal{S}' only. \square

In fact, we can go much farther in lumping states into simpler equivalence classes. Namely, let's introduce a generic state \mathbf{s} of the deletion process as $\mathbf{s} = (\nu, \nu_i, \nu_o, \mu)$, formed by the cardinalities of the three set components of \mathcal{S} : $\nu = |V|$, $\nu_i = |\mathcal{O}_i|$, and $\nu_o = |\mathcal{O}_o|$.

Now we can prove that the reduced deletion process $\{\mathbf{s}(t)\}$, formed by the cardinalities of the three set components of $\mathcal{S}(t)$ (i.e. $\mathbf{s}(t) = (|V(t)|, |\mathcal{O}_i(t)|, |\mathcal{O}_o(t)|, \mu(t))$), is Markovian as well. To this end, we will need to introduce $g(\mathbf{s}) = g(\nu, \nu_i, \nu_o, \mu)$, the number of digraphs with vertex set $[\nu]$ and μ arcs such that $\{1, \dots, \nu_i\}$ are the vertices of in-degree zero and $\{\nu_i + 1, \dots, \nu_i + \nu_o\}$ are the vertices of out-degree zero. Note that the number of digraphs, $|\mathcal{D}_{\mathcal{S}}|$, with parameter \mathcal{S} depends only upon the cardinalities of the entries of \mathcal{S} (i.e. \mathbf{s}).

Lemma 2.4. $\{\mathbf{s}(t)\}$ is a Markov process.

Proof. Let $|\mathcal{S}|$ denote $(|V|, |\mathcal{O}_i|, |\mathcal{O}_o|, \mu)$. Combining (2.1) and (2.2), we have that

$$\begin{aligned} P(\mathcal{S}(t+1) = \mathcal{S}' | \mathcal{S}(0), \dots, \mathcal{S}(t)) &= \sum_{D' \in \mathcal{D}_{\mathcal{S}'}} \left(\frac{N_{\mathcal{S},\mathcal{S}'}^{(in)}}{|\mathcal{O}_i| + |\mathcal{O}_o|} + \frac{N_{\mathcal{S},\mathcal{S}'}^{(out)}}{|\mathcal{O}_i| + |\mathcal{O}_o|} \right) \frac{1}{g(|\mathcal{S}|)} \\ &= \left(\frac{N_{\mathcal{S},\mathcal{S}'}^{(in)}}{|\mathcal{O}_i| + |\mathcal{O}_o|} + \frac{N_{\mathcal{S},\mathcal{S}'}^{(out)}}{|\mathcal{O}_i| + |\mathcal{O}_o|} \right) \frac{g(|\mathcal{S}'|)}{g(|\mathcal{S}|)}. \end{aligned}$$

Hence, if $\mathbf{s} = |\mathcal{S}|$, then

$$\begin{aligned} P(\mathbf{s}' | \mathcal{S}(t)) &:= P(\mathbf{s}(t+1) = \mathbf{s}' | \mathcal{S}(t)) = \sum_{\mathbf{s}': \mathbf{s}(\mathcal{S}') = \mathbf{s}'} \frac{N_{\mathcal{S},\mathcal{S}'}^{(in)}}{\nu_i + \nu_o} \frac{g(\mathbf{s}')}{g(\mathbf{s})} + \sum_{\mathbf{s}': \mathbf{s}(\mathcal{S}') = \mathbf{s}'} \frac{N_{\mathcal{S},\mathcal{S}'}^{(out)}}{\nu_i + \nu_o} \frac{g(\mathbf{s}')}{g(\mathbf{s})} \\ &= P_i(\mathbf{s}' | \mathcal{S}(t)) + P_o(\mathbf{s}' | \mathcal{S}(t)), \end{aligned}$$

where

$$P_i(\mathbf{s}'|\mathcal{S}(t)) := \sum_{\mathcal{S}': \mathbf{s}(\mathcal{S}') = \mathbf{s}'} \frac{N_{\mathcal{S}, \mathcal{S}'}^{(in)} g(\mathbf{s}')}{\nu_i + \nu_o g(\mathbf{s})}, \quad P_o(\mathbf{s}'|\mathcal{S}(t)) := \sum_{\mathcal{S}': \mathbf{s}(\mathcal{S}') = \mathbf{s}'} \frac{N_{\mathcal{S}, \mathcal{S}'}^{(out)} g(\mathbf{s}')}{\nu_i + \nu_o g(\mathbf{s})}.$$

Consider $P_i(\mathbf{s}'|\mathcal{S}(t))$. Recall that $N_{\mathcal{S}, \mathcal{S}'}^{(in)} = 0$ unless

$$\nu - \nu' = 1 + b, \quad \nu'_i - \nu_i = r_i - 1, \quad \nu'_o = \nu_o - b, \quad \mu' = \mu - k,$$

for some $b, r \geq 0$ and $k \geq \max\{1, b + r_i\}$; in particular, note that $|A| = a, |B| = b, |R_i| = r_i, |R_o| = r_o$. We can solve for b and r_i here in terms of \mathbf{s} and \mathbf{s}' . For any \mathcal{S} and \mathcal{S}' , so that $N_{\mathcal{S}, \mathcal{S}'}^{(in)} \neq 0$,

$$N_{\mathcal{S}, \mathcal{S}'}^{(in)} = \binom{\nu - \nu_i - b - r_i}{k - b - r_i}.$$

We wish to find the number of \mathcal{S}' with $|\mathcal{S}'| = \mathbf{s}'$ so that $N_{\mathcal{S}, \mathcal{S}'}^{(in)} \neq 0$. Any such \mathcal{S}' can be found by choosing exactly 1 vertex from \mathcal{O}_i to be deleted in Substep 1, choosing b vertices to be deleted out of \mathcal{O}_o to obtain \mathcal{O}'_o (i.e. choosing the set B), and choosing r_i vertices out of $V \setminus (\mathcal{O}_i \cup \mathcal{O}_o)$ that are added to obtain \mathcal{O}'_i (i.e. choosing the set R_i). This yields

$$|\{\mathcal{S}' : \mathbf{s}(\mathcal{S}') = \mathbf{s}', N_{\mathcal{S}, \mathcal{S}'}^{(in)} \neq 0\}| = \binom{\nu_i}{1} \binom{\nu_o}{b} \binom{\nu - \nu_i - \nu_o}{r_i}.$$

So, we find that

$$P_i(\mathbf{s}'|\mathcal{S}(t)) = \frac{1}{\nu_i + \nu_o} \frac{g(\mathbf{s}')}{g(\mathbf{s})} \binom{\nu - \nu_i - b - r_i}{k - b - r_i} \binom{\nu_i}{1} \binom{\nu_o}{b} \binom{\nu - \nu_i - \nu_o}{r_i},$$

and likewise,

$$P_o(\mathbf{s}'|\mathcal{S}(t)) = \frac{1}{\nu_i + \nu_o} \frac{g(\mathbf{s}')}{g(\mathbf{s})} \binom{\nu - \nu_i - a - r_o}{k - a - r_o} \binom{\nu_o}{1} \binom{\nu_i}{a} \binom{\nu - \nu_i - \nu_o}{r_o}.$$

We see that both $P_i(\mathbf{s}'|\mathcal{S}(t))$ and $P_o(\mathbf{s}'|\mathcal{S}(t))$ depend on $\mathcal{S}(t)$ only through $\mathbf{s}(t)$. Therefore $\{\mathbf{s}(t)\}$ is a Markov process, with the transition probability given by

$$\begin{aligned} P(\mathbf{s}'|\mathbf{s}) &= P_i(\mathbf{s}'|\mathbf{s}) + P_o(\mathbf{s}'|\mathbf{s}), \\ P_i(\mathbf{s}'|\mathbf{s}) &:= \frac{1}{\nu_i + \nu_o} \frac{g(\mathbf{s}')}{g(\mathbf{s})} \binom{\nu - \nu_i - b - r_i}{k - b - r_i} \binom{\nu_i}{1} \binom{\nu_o}{b} \binom{\nu - \nu_i - \nu_o}{r_i}, \\ P_o(\mathbf{s}'|\mathbf{s}) &:= \frac{1}{\nu_i + \nu_o} \frac{g(\mathbf{s}')}{g(\mathbf{s})} \binom{\nu - \nu_i - a - r_o}{k - a - r_o} \binom{\nu_o}{1} \binom{\nu_i}{a} \binom{\nu - \nu_i - \nu_o}{r_o}. \end{aligned} \tag{2.3}$$

The fivesome (a, b, r_i, r_o, k) is uniquely determined by the *four* components of $\Delta \mathbf{s} := (\mathbf{s}' - \mathbf{s})$ for both $P_i(\mathbf{s}'|\mathbf{s})$ and $P_o(\mathbf{s}'|\mathbf{s})$, since $a = 1, r_o = 0$ for i -transition from \mathbf{s} to \mathbf{s}' , and so

$$\Delta \mathbf{s}^T = (-1 - b, r_i - 1, -b, -k),$$

and, for o -transition, we have $b = 1, r_i = 0$, so that

$$\Delta \mathbf{s}^T = (-1 - a, -a, r_o - 1, -k).$$

With this correspondence in mind, we will denote $P_{i,o}(\mathbf{s}'|\mathbf{s}) = P_{i,o}(\Delta \mathbf{s}|\mathbf{s})$, and $P(\mathbf{s}'|\mathbf{s}) = P_i(\Delta \mathbf{s}|\mathbf{s}) + P_o(\Delta \mathbf{s}|\mathbf{s})$. \square

3 Asymptotic Transition Probability

The transition probability formulas (2.3) contain the ratio $g(\mathbf{s}')/g(\mathbf{s})$ of the counts of digraphs with parameters \mathbf{s}' and \mathbf{s} . No exact formula for those counts is available, and so our next step is to evaluate these counts sharply for \mathbf{s}' and \mathbf{s} likely to be encountered in the deletion process. To obtain a usable approximation of the transition probabilities, we will also have to approximate the binomials in those formulas, but this is considerably easier.

3.1 Counting digraphs with constrained degree sequences

To estimate $g(\mathbf{s})$, we use an argument resembling that in Pittel [32] for estimating and upper-bounding the total number of sparse digraphs with given numbers of vertices, arcs, vertices of *out*-degree zero, and no vertices of *in*-degree zero. We will need the following theorem counting the number of directed graphs with a specific in/out-degree sequence which is an important case of McKay's asymptotic formula [25], [26] for the number of $(0, 1)$ -matrices with specified row and column sums. (For very recent progress see also Greenhill and McKay [14], [15].)

Theorem 3.1. *Let $\boldsymbol{\delta} := (\delta_1, \delta_2, \dots, \delta_\nu) \geq \mathbf{0}$ and $\boldsymbol{\Delta} := (\Delta_1, \dots, \Delta_\nu) \geq \mathbf{0}$ be such that*

$$\sum_i \delta_i = \mu, \quad \sum_i \Delta_i = \mu,$$

where $\mu \geq \nu$. Introduce $g(\boldsymbol{\delta}, \boldsymbol{\Delta})$, the number of digraphs on $[\nu]$ with μ arcs with in-degree sequence $\boldsymbol{\delta}$ and out-degree sequence $\boldsymbol{\Delta}$. If $D := \max_i \delta_i + \max_i \Delta_i = O(\mu^{1/4})$, then

$$g(\boldsymbol{\delta}, \boldsymbol{\Delta}) = \frac{\mu!}{\prod_i \delta_i! \Delta_i!} H(\boldsymbol{\delta}, \boldsymbol{\Delta}),$$

where the “fudge factor”, always 1 at most, is given by

$$H(\boldsymbol{\delta}, \boldsymbol{\Delta}) = \exp \left(-\frac{1}{\mu} \sum_i \delta_i \Delta_i - \frac{1}{2\mu^2} \sum_i (\delta_i)_2 \sum_j (\Delta_j)_2 + O(D^4/\mu) \right).$$

McKay proved this theorem using a random matching scheme. First, one starts with two copies, $[\mu]_1$ and $[\mu]_2$, of $[\mu]$ along with partitions $[\mu]_1 = \cup_i I_i$, $[\mu]_2 = \cup_i O_i$, where $|I_i| = \delta_i$ and $|O_i| = \Delta_i$. Each of the $\mu!$ matchings from $[\mu]_1$ to $[\mu]_2$ gives rise to a directed multigraph (where multiple arcs between pairs of vertices and loops are allowed) where we collapse each I_i and O_i to a single vertex, v_i , and keep all present arcs.

By construction, this directed multigraph has in-degree sequence $\boldsymbol{\delta}$ and out-degree sequence $\boldsymbol{\Delta}$. Each digraph, without loops or multiple arcs, corresponds to exactly $\prod_i (\delta_i! \Delta_i!)$ matchings. So $H(\boldsymbol{\delta}, \boldsymbol{\Delta})$ is precisely the probability that a matching chosen uniformly at random among all possible matchings gives rise to a digraph without loops and multiple arcs. Thus we necessarily have $H(\boldsymbol{\delta}, \boldsymbol{\Delta}) \leq 1$, yielding

$$g(\boldsymbol{\delta}, \boldsymbol{\Delta}) \leq \frac{\mu!}{\prod_i (\delta_i! \Delta_i!)}.$$

Lemma 3.2. For all $\mathbf{s} = (\nu, \nu_i, \nu_o, \mu)$,

$$g(\mathbf{s}) \leq \mu! \frac{(e^x - 1)^{\nu - \nu_i}}{x^\mu} \frac{(e^y - 1)^{\nu - \nu_o}}{y^\mu}, \quad \forall x, y > 0. \quad (3.1)$$

Proof. First, by the definition of $g(\nu, \nu_i, \nu_o, \mu)$,

$$g(\nu, \nu_i, \nu_o, \mu) = \sum_{(\boldsymbol{\delta}, \boldsymbol{\Delta}) \in \mathbf{D}} g(\boldsymbol{\delta}, \boldsymbol{\Delta}),$$

where \mathbf{D} is the set of all pairs of admissible in-degree sequences and out-degree sequences for the 4-tuple (ν, ν_i, ν_o, μ) . Formally, \mathbf{D} is defined by constraints

$$\sum_{j=1}^{\nu} \delta_j = \sum_{j=1}^{\nu} \Delta_j = \mu, \quad (3.2)$$

and

$$\begin{aligned} \delta_j &= 0, \text{ for } j \in \{1, \dots, \nu_i\}, \\ \delta_j &> 0 \text{ for } j > \nu_i, \\ \Delta_j &= 0 \text{ for } j \in \{\nu_i + 1, \dots, \nu_i + \nu_o\}, \\ \Delta_j &> 0 \text{ for } j \notin \{\nu_i + 1, \dots, \nu_i + \nu_o\}. \end{aligned} \quad (3.3)$$

The constraint (3.2) calls for using the bivariate generating function of the counts of pairs $(\boldsymbol{\delta}, \boldsymbol{\Delta})$ by the values of $\sum_j \delta_j$ and $\sum_j \Delta_j$. From McKay's Formula (Theorem 3.1), we see that

$$\begin{aligned} g(\mathbf{s}) &= \mu! [x^\mu y^\mu] \sum_{(\boldsymbol{\delta}, \boldsymbol{\Delta}) \in \mathbf{D}} H(\boldsymbol{\delta}, \boldsymbol{\Delta}) \prod_j \frac{x^{\delta_j} y^{\Delta_j}}{\delta_j! \Delta_j!} \\ &\leq \mu! [x^\mu y^\mu] \sum_{(\boldsymbol{\delta}, \boldsymbol{\Delta}) \text{ meets (3.3)}} \prod_j \frac{x^{\delta_j}}{\delta_j!} \cdot \frac{y^{\Delta_j}}{\Delta_j!}. \end{aligned} \quad (3.4)$$

Since the constraints (3.3) are imposed only on individual δ_j, Δ_j , the last sum can be factored into the product of simple series:

$$\begin{aligned} g(\mathbf{s}) &\leq \mu! [x^\mu y^\mu] \prod_{j=1}^{\nu_i} \sum_{\Delta_j \geq 1} \frac{y^{\Delta_j}}{\Delta_j!} \cdot \prod_{j=\nu_i+1}^{\nu_i+\nu_o} \sum_{\delta_j \geq 1} \frac{x^{\delta_j}}{\delta_j!} \cdot \prod_{j=\nu_i+\nu_o+1}^{\nu} \sum_{\delta_j, \Delta_j \geq 1} \frac{x^{\delta_j} y^{\Delta_j}}{\delta_j! \Delta_j!} \\ &= \mu! [x^\mu y^\mu] ((e^y - 1)^{\nu_i} \cdot (e^x - 1)^{\nu_o} \cdot ((e^x - 1)(e^y - 1))^{\nu - \nu_i - \nu_o}) \\ &\leq \mu! x^{-\mu} y^{-\mu} ((e^y - 1)^{\nu_i} \cdot (e^x - 1)^{\nu_o} \cdot ((e^x - 1)(e^y - 1))^{\nu - \nu_i - \nu_o}), \end{aligned}$$

for all $x, y > 0$. □

Naturally, we wish to determine the values of x and y which minimize the RHS in (3.1).

Lemma 3.3. Suppose $\mu > \nu - \nu_i$, $\nu - \nu_o > 0$. If

$$\phi_i(x) := \frac{(e^x - 1)^{\nu - \nu_i}}{x^\mu}, \quad \phi_o(y) := \frac{(e^y - 1)^{\nu - \nu_o}}{y^\mu},$$

then $\phi_i(x)\phi_o(y)$ is minimized at $x = z_i, y = z_o$, where z_i, z_o are the unique positive roots of

$$\frac{z_i e^{z_i}}{e^{z_i} - 1} = \frac{\mu}{\nu - \nu_i}, \quad \frac{z_o e^{z_o}}{e^{z_o} - 1} = \frac{\mu}{\nu - \nu_o}.$$

Remark 3.4. We shall see later that the non-zero in-degree (the non-zero out-degree, resp.) of a generic vertex in a digraph chosen uniformly at random among all $g(\mathbf{s})$ digraphs is close in distribution to a Poisson $Z(z_i)$ ($Z(z_o)$, resp.) conditioned on $Z(z_i) \geq 1$ ($Z(z_o) \geq 1$, resp.). In fact, these variables z_i and z_o are important for the asymptotics of $g(\mathbf{s})$. Such seemingly “hidden” parameters turned out to be ubiquitous in situations requiring asymptotic enumeration of graphs with constrained degree sequences; see, [1], [36], [33].

Proof. Since ϕ_i and ϕ_o are positive functions, it suffices to minimize $\phi_i(x)$ and $\phi_o(y)$ separately. Consider $\phi_i(x)$. We find that

$$\frac{d}{dx} \ln \phi_i(x) = (\nu - \nu_i) \frac{e^x}{e^x - 1} - \frac{\mu}{x} = \frac{\nu - \nu_i}{x} \left(\frac{x e^x}{e^x - 1} - \frac{\mu}{\nu - \nu_i} \right).$$

The function $\ell(x) := \frac{x e^x}{e^x - 1}$ is strictly increasing on $(0, \infty)$, where $\ell(0+) = 1 < \frac{\mu}{\nu - \nu_i}$ and $\ell(\infty) = \infty$. Hence, $\phi_i(x)$ does attain its minimum at the unique $z_i = z_i(\mathbf{s})$, satisfying $\ell(z_i) = \frac{\mu}{\nu - \nu_i}$. Likewise $\phi_o(y)$ attains its minimum at the unique $z_o = z_o(\mathbf{s})$, satisfying $\ell(z_o) = \frac{\mu}{\nu - \nu_o}$. \square

Now we can sharply approximate $g(\mathbf{s})$ for a good range of \mathbf{s} .

Theorem 3.5. Let $\mathbf{s} = (\nu, \nu_i, \nu_o, \mu)$ be such that $\nu \rightarrow \infty$, $\mu - \nu \rightarrow \infty$, $\mu = O(\nu)$ and $\nu - \nu_i - \nu_o = \Theta(\nu)$. Introduce Z^i, Z^o , two independent truncated Poissons, with parameters $z_i = z_i(\mathbf{s})$ and $z_o = z_o(\mathbf{s})$ from Lemma 3.3, i.e.

$$P(Z^i = j) = \frac{z_i^j / j!}{e^{z_i} - 1}, \quad P(Z^o = j) = \frac{z_o^j / j!}{e^{z_o} - 1}, \quad (j \geq 1).$$

Then

$$\begin{aligned} g(\mathbf{s}) &= \left(1 + O\left(\frac{(\ln \nu)^6}{\nu}\right) \right) \mu! \frac{(e^{z_i} - 1)^{\nu - \nu_i} (e^{z_o} - 1)^{\nu - \nu_o}}{z_i^\mu z_o^\mu} \\ &\quad \times \frac{e^{-\eta}}{2\pi \sqrt{(\nu - \nu_i) \text{Var}[Z^i] (\nu - \nu_o) \text{Var}[Z^o]}}, \end{aligned} \tag{3.5}$$

where

$$\eta = \frac{\mu(\nu - \nu_i - \nu_o)}{(\nu - \nu_i)(\nu - \nu_o)} + \frac{z_i z_o}{2}.$$

This theorem and its proof are similar to Theorem 2.2 in [32].

Proof. First rewrite (3.4) as

$$g(\nu, \nu_i, \nu_o, \mu) = \mu! \frac{(e^{z_i} - 1)^{\nu - \nu_i} (e^{z_o} - 1)^{\nu - \nu_o}}{(z_i z_o)^\mu} \times \sum_{(\boldsymbol{\delta}, \boldsymbol{\Delta}) \in \mathcal{D}} H(\boldsymbol{\delta}, \boldsymbol{\Delta}) [x^\mu y^\mu] \prod_{j: \delta_j > 0} \frac{(z_i x)^{\delta_j} / \delta_j!}{e^{z_i} - 1} \cdot \prod_{j: \Delta_j > 0} \frac{(z_o y)^{\Delta_j} / \Delta_j!}{e^{z_o} - 1}. \quad (3.6)$$

Further we notice that $\frac{(z_i)^{\delta_j} / \delta_j!}{e^{z_i} - 1}$ is precisely the probability that Z^i equals δ_j ; $\frac{(z_o)^{\Delta_j} / \Delta_j!}{e^{z_o} - 1}$ is interpreted similarly. The products over j in (3.6) force us to introduce the sequences of independent copies of Z^i and Z^o . Define $Z_{\nu_i+1}^i, Z_{\nu_i+2}^i, \dots, Z_\nu^i$ as independent copies of Z^i , and $Z_1^o, \dots, Z_{\nu_i}^o, Z_{\nu_i+\nu_o+1}^o, \dots, Z_\nu^o$ as independent copies of Z^o , and introduce

$$\begin{aligned} \mathbf{Z}^i &= (\nu_i \text{ zeroes}, Z_{\nu_i+1}^i, \dots, Z_\nu^i), \\ \mathbf{Z}^o &= (Z_1^o, \dots, Z_{\nu_i}^o, \nu_o \text{ zeroes}, Z_{\nu_i+\nu_o+1}^o, \dots, Z_\nu^o). \end{aligned}$$

For the coordinates that are zero, we define $Z_{j'}^i$ and $Z_{j''}^o$ to be zero. Because all $Z_{j'}^i, Z_{j''}^o$ are independent, the factor by $H(\boldsymbol{\delta}, \boldsymbol{\Delta})$ in the second line of (3.6) is $P(\mathbf{Z}^i = \boldsymbol{\delta}, \mathbf{Z}^o = \boldsymbol{\Delta})$. Therefore the expression in this line is the expected value of $H(\mathbf{Z}^i, \mathbf{Z}^o)$, *conditioned* on the event $\{\|\mathbf{Z}^i\| = \mu, \|\mathbf{Z}^o\| = \mu\}$; ($\|\{x_j\}\| \stackrel{\text{def}}{=} \sum_j |x_j|$, the 1-norm of $\{x_j\}$). To get a handle on this conditional expectation, we notice first that

$$\begin{aligned} E[H(\mathbf{Z}^i, \mathbf{Z}^o) 1_{\{\|\mathbf{Z}^i\| = \|\mathbf{Z}^o\| = \mu\}}] &= E\left[H(\mathbf{Z}^i, \mathbf{Z}^o) \middle| \|\mathbf{Z}^i\| = \|\mathbf{Z}^o\| = \mu\right] \\ &\quad \times P(\|\mathbf{Z}^i\| = \mu) P(\|\mathbf{Z}^o\| = \mu). \end{aligned}$$

Furthermore, by a local limit theorem, (see [1], Appendix), one can show that, under conditions of Theorem 3.5,

$$P(\|\mathbf{Z}^i\| = \mu) = \frac{1 + O(1/\nu)}{\sqrt{2\pi(\nu - \nu_i) \text{Var}[Z^i]}}, \quad P(\|\mathbf{Z}^o\| = \mu) = \frac{1 + O(1/\nu)}{\sqrt{2\pi(\nu - \nu_o) \text{Var}[Z^o]}}.$$

And just like [32], (see also Pittel and Wormald [35]), one can show that this conditional expectation of $H(\mathbf{Z}^i, \mathbf{Z}^o)$ is within a multiplicative factor $(1 + O((\ln \nu)^6 / \nu))$ from

$$\exp \left\{ -\frac{1}{\mu} E \left[\sum_j Z_j^i Z_j^o \right] - \frac{1}{2} \left(\frac{1}{\mu} E \left[\sum_j (Z_j^i)_2 \right] \right) \left(\frac{1}{\mu} E \left[\sum_j (Z_j^o)_2 \right] \right) \right\},$$

where $(Z)_2 := Z(Z-1)$. Since Z_j^i and Z_j^o are independent, we have $E[Z_j^i Z_j^o] = E[Z_j^i] E[Z_j^o]$. For the first $\nu_i + \nu_o$ coordinates, either $E[Z_j^i] = 0$ or $E[Z_j^o] = 0$. For the last $\nu - \nu_i - \nu_o$ entries,

$$E[Z_j^i] = E[Z^i] = \frac{z_i}{1 - e^{-z_i}} = \frac{\mu}{\nu - \nu_i}$$

and

$$E[Z_j^o] = E[Z^o] = \frac{\mu}{\nu - \nu_o}.$$

Hence

$$\frac{1}{\mu} E \left[\sum_j Z_j^i Z_j^o \right] = \frac{\mu(\nu - \nu_i - \nu_o)}{(\nu - \nu_i)(\nu - \nu_o)}.$$

Furthermore, if Z_j^i is non-zero, then

$$E[(Z_j^i)_2] = E[(Z^i)_2] = z_i^2 / (1 - e^{-z_i}) = z_i \mu / (\nu - \nu_i).$$

Since exactly $\nu - \nu_i$ of the Z_j^i are non-zero, we have that

$$\frac{1}{\mu} E \left[\sum_j (Z_j^i)_2 \right] = \frac{\nu - \nu_i}{\mu} \cdot \frac{z_i \mu}{\nu - \nu_i} = z_i.$$

Likewise, we have

$$\frac{1}{\mu} E \left[\sum_j (Z_j^o)_2 \right] = z_o$$

which completes the proof of the theorem. \square

Since the asymptotic count of digraphs with parameter \mathbf{s} is, in essence, an approximate estimate of the number of most numerous digraphs, the following interpretation of Z^i and Z^o is hardly surprising. Pick an admissible digraph uniformly at random among all $g(\mathbf{s})$ such digraphs. Then, under conditions of Theorem 3.5, the out-degree (in-degree) of a fixed vertex in the set $\{\nu_i + 1, \dots, n\}$ ($\{1, \dots, \nu_i\} \cup \{\nu_i + \nu_o + 1, \dots, n\}$ resp.) is asymptotic, in distribution, to Z^o (Z^i resp.). Moreover, for a fixed set of vertices, their in/out-degrees are asymptotically independent.

3.2 Approximating transition probabilities

Now that we have found an asymptotic formula for $g(\mathbf{s})$, we can find asymptotic formulas for our transition probabilities defined in (2.3). To do so, we will impose stricter conditions on \mathbf{s} . First define two functions,

$$F_1(\mathbf{s}) := \frac{\mu(\nu - \nu_i - \nu_o)}{(\nu - \nu_i)(\nu - \nu_o)}, \quad F_2(\mathbf{s}) := \frac{z_i(\mathbf{s})z_o(\mathbf{s})}{\mu/n}. \quad (3.7)$$

Implicit in this definition is the constraint on \mathbf{s} :

$$\mu, \nu - \nu_i, \nu - \nu_o > 0, \quad \nu \leq n, \quad \mu \leq c_n n, \quad \text{and} \quad \frac{\mu}{\nu - \nu_i}, \frac{\mu}{\nu - \nu_o} > 1. \quad (3.8)$$

We will eventually prove that, for the initial states \mathbf{s}_0 in the likely range arising from $D(n, m = c_n n)$, the values of F_1 and F_2 a.a.s. are almost constant throughout the deletion process. That's what makes $F_i(\mathbf{s})$ instrumental in our analysis of the deletion algorithm.

Remark 3.6. Notice that both functions appear in the exponent $\eta = \eta(\mathbf{s})$ in Theorem 3.5. Their appearance is coincidental, as far as we can tell.

Since $c_n \rightarrow c \in (1, \infty)$, we may and will assume that c_n is bounded away from 1. Suppose that $\epsilon = \epsilon(n) > 0$ is such that $c_n - \epsilon \geq 1 + \delta$, for some fixed $\delta > 0$. Define

$$\mathbf{S}_\epsilon := \{\mathbf{s} : \mathbf{s} \text{ meets (3.8); } F_1(\mathbf{s}), F_2(\mathbf{s}) \in (c_n - \epsilon, c_n + \epsilon); \nu_i + \nu_o > 0\}. \quad (3.9)$$

Fact 3.7. *Uniformly over $\mathbf{s} \in \mathbf{S}_\epsilon$,*

- (i) $\nu, \nu - \nu_i, \nu - \nu_o, \nu - \nu_i - \nu_o = \Theta(n)$;
- (ii) $\mu - \nu = \Theta(n)$;
- (iii) $z_i(\mathbf{s})$ and $z_o(\mathbf{s})$ (defined in Lemma 3.3) are bounded away from 0 and ∞ .

In particular, the conditions of Theorem 3.5 are met uniformly for $\mathbf{s} \in \mathbf{S}_\epsilon$.

Proof. Since $F_1(\mathbf{s}) > 0$, we have $\nu - \nu_i - \nu_o > 0$. That $z_i(\mathbf{s})$ and $z_o(\mathbf{s})$ are bounded away from 0 follows from

$$\frac{\mu}{\nu - \nu_i}, \frac{\mu}{\nu - \nu_o} \geq F_1(\mathbf{s}) \geq c_n - \epsilon > 1.$$

Now, because $z_i(\mathbf{s})$ and $z_o(\mathbf{s})$ are bounded away from 0 and $z_i(\mathbf{s})z_o(\mathbf{s})/(\mu/n) = F_2(\mathbf{s}) \leq c_n + \epsilon$, we have that μ/n is bounded away from zero as well. In addition, since $\mu = O(n)$, $F_2(\mathbf{s}) \leq c_n + \epsilon$ and $z_i(\mathbf{s})$ is bounded away from zero, we conclude that $z_o(\mathbf{s})$ is bounded away from ∞ ; likewise $z_i(\mathbf{s})$ is bounded away from ∞ as well. And this implies the part (i). Finally, the part (ii) follows then from

$$\mu = F_1(\mathbf{s}) \frac{(\nu - \nu_i)(\nu - \nu_o)}{\nu - \nu_i - \nu_o} = F_1(\mathbf{s}) \left[\nu + \frac{\nu_i \nu_o}{\nu - \nu_i - \nu_o} \right],$$

and the condition $F_1(\mathbf{s}) \geq c_n - \epsilon$. □

With this preliminaries done, we focus on the factor $g(\mathbf{s}')/g(\mathbf{s})$ in the formulas (2.3) for the transition probabilities $P_i(\Delta \mathbf{s}|\mathbf{s})$ and $P_o(\Delta \mathbf{s}|\mathbf{s})$. In the next statement and in the rest of the paper we will use a notation $A \lesssim B$, meaning $A = O(B)$, when expression for B is too bulky.

Lemma 3.8. *Let $\mathbf{s}' = \mathbf{s} + \Delta \mathbf{s}$ and $\Delta \mathbf{s} = (-a - b, r_i - a, r_o - b, -k)$. Uniformly over $\mathbf{s} \in \mathbf{S}_\epsilon$,*

(i)
$$\frac{g(\mathbf{s}')}{g(\mathbf{s})} \lesssim \frac{\nu}{(\mu)_k} \frac{(z_i z_o)^k}{(e^{z_i} - 1)^{b+r_i} (e^{z_o} - 1)^{a+r_o}};$$

(ii) *if, in addition, $a = 1, r_o = 0$ and $\max\{1, b + r_i\} \leq k \leq \ln n$, then*

$$\frac{g(\mathbf{s}')}{g(\mathbf{s})} = \left(1 + O\left(\frac{(\ln n)^6}{n}\right) \right) \left(\frac{z_i z_o}{\mu} \right)^k \frac{1}{(e^{z_o} - 1)(e^{z_i} - 1)^{b+r_i}};$$

likewise, if $b = 1, r_i = 0$ and $\max\{1, a + r_o\} \leq k \leq \ln n$, then

$$\frac{g(\mathbf{s}')}{g(\mathbf{s})} = \left(1 + O\left(\frac{(\ln n)^6}{n}\right) \right) \left(\frac{z_i z_o}{\mu} \right)^k \frac{1}{(e^{z_i} - 1)(e^{z_o} - 1)^{a+r_o}}.$$

Proof. (i) First, by Theorem 3.5, we have that

$$g(\mathbf{s}) = \left(1 + O\left(\frac{(\ln n)^6}{n}\right)\right) \frac{\mu! h_i(\mathbf{s}, z_i) h_o(\mathbf{s}, z_o) e^{-\eta(\mathbf{s})}}{2\pi \sqrt{(\nu - \nu_i) \text{Var}[Z^i] (\nu - \nu_o) \text{Var}[Z^o]}},$$

where

$$h_i(\mathbf{s}, z_i) = \frac{(e^{z_i} - 1)^{\nu - \nu_i}}{z_i^\mu}, \quad h_o(\mathbf{s}, z_o) = \frac{(e^{z_o} - 1)^{\nu - \nu_o}}{z_o^\mu},$$

and

$$\eta(\mathbf{s}) = F_1(\mathbf{s}) + F_2(\mathbf{s}) \frac{\mu}{2\nu}.$$

Uniformly over $\mathbf{s} \in \mathbf{S}_\epsilon$, $\eta(\mathbf{s})$, $z_i(\mathbf{s})$ and $z_o(\mathbf{s})$ are bounded away from ∞ . So the variances of Z^i and Z^o are bounded as well. Hence

$$\frac{1}{g(\mathbf{s})} \lesssim \frac{\nu}{\mu!} \frac{z_i^\mu}{(e^{z_i} - 1)^{\nu - \nu_i}} \frac{z_o^\mu}{(e^{z_o} - 1)^{\nu - \nu_o}}.$$

From Lemma 3.2 using $x = z_i(\mathbf{s})$ and $y = z_o(\mathbf{s})$, we also have that

$$g(\mathbf{s}') \leq (\mu - k)! \frac{(e^{z_i} - 1)^{\nu - \nu_i - b - r_i}}{z_i^{\mu - k}} \frac{(e^{z_o} - 1)^{\nu - \nu_o - a - r_o}}{z_o^{\mu - k}},$$

where, crucially, $z_i = z_i(\mathbf{s})$ and $z_o = z_o(\mathbf{s})$ rather than $z_i(\mathbf{s}')$ and $z_o(\mathbf{s}')$. Consequently

$$\frac{g(\mathbf{s}')}{g(\mathbf{s})} \lesssim \frac{\nu}{(\mu)_k} \frac{(z_i z_o)^k}{(e^{z_i} - 1)^{b + r_i} (e^{z_o} - 1)^{a + r_o}}.$$

(ii) Consider, for instance, the case $a = 1$, $r_o = 0$ and $\max\{1, b + r_i\} \leq k \leq \ln n$. Here \mathbf{s}' also meets the conditions of Theorem 3.5, and so we have the asymptotic formula (3.5) for $g(\mathbf{s}')$. To estimate sharply $g(\mathbf{s}')/g(\mathbf{s})$, we need to look closely at the difference between $z_{i,o}(\mathbf{s})$ and $z_{i,o}(\mathbf{s}')$, and between $h_{i,o}(\mathbf{s}, z_{i,o}(\mathbf{s}))$ and $h_{i,o}(\mathbf{s}', z_{i,o}(\mathbf{s}'))$.

First of all,

$$\nu' = \nu + O(\ln n), \quad \nu'_i = \nu_i + O(\ln n), \quad \nu'_o = \nu_o + O(\ln n), \quad \mu' = \mu + O(\ln n).$$

Then, denoting $z'_{i,o} = z_{i,o}(\mathbf{s}')$, we have that

$$\ell(z'_i) = \frac{\mu'}{\nu' - \nu'_i} = \frac{\mu - k}{\nu - \nu_i - b - r_i} = \frac{\mu}{\nu - \nu_i} + O((\ln n)/n) = \ell(z_i) + O((\ln n)/n),$$

where as before $\ell(z) = \frac{ze^z}{e^z - 1}$. Since $\ell'(z) \in [1/2, 1]$ for $z \geq 0$, it follows then that $z'_i = z_i + O((\ln n)/n)$, and similarly, $z'_o = z_o + O((\ln n)/n)$. Therefore, denoting $(Z^{i,o})' = Z^{i,o}(\mathbf{s}')$, $\eta' = \eta(\mathbf{s}')$,

$$\text{Var}[(Z^i)'] = \text{Var}[Z^i](1 + O((\ln n)/n)), \quad \text{Var}[(Z^o)'] = \text{Var}[Z^o](1 + O((\ln n)/n)),$$

and

$$\eta' = \eta + O((\ln n)/n).$$

Consequently,

$$\frac{g(\mathbf{s}')}{g(\mathbf{s})} = (1 + O((\ln n)^6/n)) \frac{1}{\mu^k} \frac{h_i(\mathbf{s}', z'_i)}{h_i(\mathbf{s}, z_i)} \cdot \frac{h_o(\mathbf{s}', z'_o)}{h_o(\mathbf{s}, z_o)}.$$

Notice that

$$\frac{h_i(\mathbf{s}', z'_i)}{h_i(\mathbf{s}, z_i)} = \frac{(e^{z'_i} - 1)^{\nu' - \nu'_i} / (z'_i)^{\mu'}}{(e^{z_i} - 1)^{\nu' - \nu'_i} / (z_i)^{\mu'}} \frac{z_i^k}{(e^{z_i} - 1)^{b+r_i}} = \frac{h_i(\mathbf{s}', z'_i)}{h_i(\mathbf{s}, z_i)} \frac{z_i^k}{(e^{z_i} - 1)^{b+r_i}}$$

and

$$\frac{h_o(\mathbf{s}', z'_o)}{h_o(\mathbf{s}, z_o)} = \frac{(e^{z'_o} - 1)^{\nu' - \nu'_o} / (z'_o)^{\mu'}}{(e^{z_o} - 1)^{\nu' - \nu'_o} / (z_o)^{\mu'}} \frac{z_o^k}{e^{z_o} - 1} = \frac{h_o(\mathbf{s}', z'_o)}{h_o(\mathbf{s}, z_o)} \frac{z_o^k}{e^{z_o} - 1}.$$

To complete the proof, it suffices to show that, uniformly over \mathbf{s} and \mathbf{s}' in question,

$$\frac{h_i(\mathbf{s}', z'_i)}{h_i(\mathbf{s}, z_i)} = (1 + O((\ln n)^2/n)), \quad \frac{h_o(\mathbf{s}', z'_o)}{h_o(\mathbf{s}, z_o)} = (1 + O((\ln n)^2/n)). \quad (3.10)$$

To this end, we expand the exponent of $h_i(\mathbf{s}', z)$ about $z = z_i$:

$$\begin{aligned} h_i(\mathbf{s}', z'_i) &= \exp \left((\nu' - \nu'_i) \ln(e^{z'_i} - 1) - \mu' \ln z'_i \right) \\ &= \exp \left((\nu' - \nu'_i) \ln(e^{z_i} - 1) - \mu' \ln z_i \right) \\ &\quad \times \exp \left(\left((\nu' - \nu'_i) \frac{e^{z_i}}{e^{z_i} - 1} - \mu' / z_i \right) (z'_i - z_i) + O(\nu \cdot (z'_i - z_i)^2) \right). \end{aligned}$$

Since $z'_i - z_i = O((\ln n)/n)$, the big-Oh term is $O((\ln n)^2/n)$. Also, by definition of z_i ,

$$\begin{aligned} (\nu' - \nu'_i) \frac{e^{z_i}}{e^{z_i} - 1} - \frac{\mu'}{z_i} &= (\nu - \nu_i) \frac{e^{z_i}}{e^{z_i} - 1} - \frac{\mu}{z_i} + O(\ln n) \\ &= \frac{\nu - \nu_i}{z_i} \left(\ell(z_i) - \frac{\mu}{\nu - \nu_i} \right) + O(\ln n) = O(\ln n). \end{aligned}$$

So the first equation in (3.10) follows. The second equation is proved similarly. \square

Besides $g(\mathbf{s}')/g(\mathbf{s})$, the factors in formulas (2.3) and for P_i and P_o are binomials. In the case of P_i , they are:

$$\binom{\nu - \nu_i - b - r_i}{k - b - r_i}, \binom{\nu - \nu_i - \nu_o}{r_i}, \text{ and } \binom{\nu_o}{b}.$$

For \mathbf{s} and $\Delta \mathbf{s}$ that meet the conditions of Lemma 3.8 (ii), we have that

$$\begin{aligned} \binom{\nu - \nu_i - b - r_i}{k - b - r_i} &= \frac{(\nu - \nu_i)^{k-b-r_i}}{(k - b - r_i)!} \prod_{i=0}^{k-b-r_i-1} \frac{\nu - \nu_i - b - r_i - i}{\nu - \nu_i} \\ &= \frac{(\nu - \nu_i)^{k-b-r_i}}{(k - b - r_i)!} (1 + O((\ln n)^2/n)), \end{aligned} \quad (3.11)$$

and similarly,

$$\binom{\nu - \nu_i - \nu_o}{r_i} = \frac{(\nu - \nu_i - \nu_o)^{r_i}}{r_i!} (1 + O((\ln n)^2/n)). \quad (3.12)$$

However, we leave $\binom{\nu_o}{b}$ as it stands, since close to the end of the deletion process, ν_o , the number of vertices with out-degree zero, can not be expected to be much larger than b , the number of out-degree zero vertices deleted in one step.

Motivated by these asymptotic expressions and the sharp asymptotics of the ratio $g(\mathbf{s}')/g(\mathbf{s})$ in Lemma 3.8, we define the following auxiliary transition “probabilities”:

$$\begin{aligned} q_i(\Delta \mathbf{s} | \mathbf{s}) &:= \frac{\nu_i}{\nu_i + \nu_o} \frac{1}{e^{z_o} - 1} \binom{\nu_o}{b} \left(\frac{1}{e^{z_i} - 1} \frac{z_i z_o}{\mu} \right)^b \frac{1}{r_i!} \left(\frac{\nu - \nu_i - \nu_o}{e^{z_i} - 1} \frac{z_i z_o}{\mu} \right)^{r_i} \\ &\times \frac{1}{(k - b - r_i)!} \left((\nu - \nu_i) \frac{z_i z_o}{\mu} \right)^{k - b - r_i}, \end{aligned} \quad (3.13)$$

if $\Delta \mathbf{s}$ is such that $a = 1, r_o = 0$ and $k \geq \max\{1, b + r_i\}$, and $q_i = 0$ for any other $\Delta \mathbf{s}$. Likewise

$$\begin{aligned} q_o(\Delta \mathbf{s} | \mathbf{s}) &:= \frac{\nu_o}{\nu_i + \nu_o} \frac{1}{e^{z_i} - 1} \binom{\nu_i}{a} \left(\frac{1}{e^{z_o} - 1} \frac{z_i z_o}{\mu} \right)^a \frac{1}{r_o!} \left(\frac{\nu - \nu_i - \nu_o}{e^{z_o} - 1} \frac{z_i z_o}{\mu} \right)^{r_o} \\ &\times \frac{1}{(k - a - r_o)!} \left((\nu - \nu_o) \frac{z_i z_o}{\mu} \right)^{k - a - r_o}, \end{aligned}$$

if $\Delta \mathbf{s}$ is such that $b = 1, r_i = 0$ and $k \geq \max\{1, a + r_o\}$, and $q_o = 0$ for any other $\Delta \mathbf{s}$. Although it is not immediately apparent, $q_i + q_o$ is a substochastic distribution, meaning that $\sum_{\Delta \mathbf{s}} [q_i(\Delta \mathbf{s} | \mathbf{s}) + q_o(\Delta \mathbf{s} | \mathbf{s})] < 1$, with the probability deficit being exponentially small. Sure enough, $q_{i,o}(\Delta \mathbf{s} | \mathbf{s})$ closely approximates $P_{i,o}(\Delta \mathbf{s} | \mathbf{s})$ for \mathbf{s} and $\Delta \mathbf{s}$ that matter.

Lemma 3.9. (i) *Uniformly over $\mathbf{s} \in \mathbf{S}_\epsilon$, defined in (3.9), and $\Delta \mathbf{s}$ with $k \leq \ln n$,*

$$\begin{aligned} q_i(\Delta \mathbf{s} | \mathbf{s}) &= P_i(\Delta \mathbf{s} | \mathbf{s}) (1 + O((\ln n)^6/n)), \\ q_o(\Delta \mathbf{s} | \mathbf{s}) &= P_o(\Delta \mathbf{s} | \mathbf{s}) (1 + O((\ln n)^6/n)). \end{aligned}$$

(ii) *Uniformly over $\mathbf{s} \in \mathbf{S}_\epsilon$,*

$$\sum_{\Delta \mathbf{s}: k \geq \ln n} q_i(\Delta \mathbf{s} | \mathbf{s}), \quad \sum_{\Delta \mathbf{s}: k \geq \ln n} P_i(\Delta \mathbf{s} | \mathbf{s}) \leq \exp \left(-\frac{2}{3} (\ln n) (\ln \ln n) \right)$$

and

$$\sum_{\Delta \mathbf{s}: k \geq \ln n} q_o(\Delta \mathbf{s} | \mathbf{s}), \quad \sum_{\Delta \mathbf{s}: k \geq \ln n} P_o(\Delta \mathbf{s} | \mathbf{s}) \leq \exp \left(-\frac{2}{3} (\ln n) (\ln \ln n) \right).$$

Proof. By symmetry, we need only consider q_i and P_i for both parts of the lemma.

(i) This equality is immediate from part (ii) of Lemma 3.8 and the binomial expression approximations (3.11) and (3.12).

(ii) First, by (2.3), we have that

$$P_i(\Delta \mathbf{s} | \mathbf{s}) \leq \frac{g(\mathbf{s}')}{g(\mathbf{s})} \binom{\nu - \nu_i - b - r_i}{k - b - r_i} \binom{\nu_o}{b} \binom{\nu - \nu_i - \nu_o}{r_i} \leq \frac{g(\mathbf{s}')}{g(\mathbf{s})} 2^k \binom{\nu - \nu_i}{k}. \quad (3.14)$$

Indeed, the product of the second and third binomials in (3.14) is at most $\binom{\nu - \nu_i}{b + r_i}$, the number of ways to sample $b + r_i$ balls out of the urn with ν_o blue balls and $\nu - \nu_i - \nu_o$ red balls. And $\binom{\nu - \nu_i}{b + r_i}$ times the first binomial in (3.14) is at most $2^k \binom{\nu - \nu_i}{k}$, the total number of ways to sample k balls out of the urn with $\nu - \nu_i$ colorless balls *and* color each of sampled balls either white or green.

So, using the bound for $g(\mathbf{s}')/g(\mathbf{s})$ from part (i) of Lemma 3.8, we have that

$$P_i(\Delta \mathbf{s} | \mathbf{s}) \leq \frac{\nu}{(\mu)_k} \frac{(2z_i z_o)^k}{(e^{z_i} - 1)^{b+r_i} (e^{z_o} - 1)} \frac{(\nu)_k}{k!} \leq \frac{\nu}{k!} \frac{(2z_i z_o)^k}{(e^{z_i} - 1)^{b+r_i} (e^{z_o} - 1)}.$$

Since $\mathbf{s} \in \mathbf{S}_\epsilon$, z_i and z_o are bounded above by some fixed $A > 0$, see Fact 3.7. Using $(e^x - 1)^{-1} < x^{-1}$, we have then that

$$P_i(\Delta \mathbf{s} | \mathbf{s}) \leq \frac{\nu}{k!} 2^k (z_i)^{k-b-r_i} (z_o)^{k-1} \leq \frac{\nu (2A^2)^k}{k!},$$

uniformly for b, r_i with $b + r_i \leq k$. Consequently

$$\sum_{b \geq 0} \sum_{r_i \geq 0} \sum_{k \geq \max\{b+r_i, \ln \nu\}} P_i(\Delta \mathbf{s} | \mathbf{s}) = O\left(\nu^2 \frac{(2A^2)^{\ln \nu}}{[\ln \nu]!}\right).$$

Using Stirling's formula on $[\ln \nu]!$ yields

$$\begin{aligned} \sum_{\Delta \mathbf{s}: k \geq \ln \nu} P_i(\Delta \mathbf{s} | \mathbf{s}) &\lesssim \nu^3 \frac{(2eA^2)^{\ln \nu}}{(\ln \nu)^{\ln \nu}} = \exp(-(\ln \nu)(\ln \ln \nu) + O(\ln \nu)) \\ &= \exp(-(\ln n)(\ln \ln n) + O(\ln n)), \end{aligned}$$

uniformly for $\mathbf{s} \in \mathbf{S}_\epsilon$. The other three inequalities can be proved in a similar fashion. \square

3.3 Estimating expectations

Lemma 3.9 allows us to sharply approximate the (conditional) expected values of all five parameters a, b, r_i, r_o, k and all 15 pairwise products of these parameters by the “expected” values with respect to the approximation $q(\Delta \mathbf{s} | \mathbf{s}) := q_i(\Delta \mathbf{s} | \mathbf{s}) + q_o(\Delta \mathbf{s} | \mathbf{s})$. These approximations will be instrumental for the key proofs. For brevity, given any function f of parameters a, b, r, t, k , we denote the corresponding expected values $E_s[f]$ and $E_s^q[f]$. To be sure, $E_s^q[f]$ is not exactly the expected value of f since $q(\Delta \mathbf{s} | \mathbf{s})$ is substochastic.

Lemma 3.10. *If f is a linear or quadratic function of the components of $\Delta \mathbf{s}$,*

$$E_{\mathbf{s}}[f] = E_{\mathbf{s}}^q[f] + O((\ln n)^8/n),$$

uniformly for $\mathbf{s} \in \mathbf{S}_\epsilon$,

Proof. Observe that $|f| = O((\ln n)^2)$ if $k \leq \ln n$, and $|f| = O(n^2)$ always. By Lemma 3.9 (ii), the contributions to $E_{\mathbf{s}}[f]$ and $E_{\mathbf{s}}^q[f]$ coming from $k \geq \ln n$ are on order at most

$$n^3 \exp\left(-\frac{2}{3} \ln n (\ln \ln n)\right) \ll \frac{(\ln n)^8}{n}.$$

Furthermore, by Lemma 3.9 (i), and the first condition on f ,

$$\begin{aligned} & \left| \sum_{\Delta \mathbf{s}: k \leq \ln n} f(a, b, r_i, r_o, k) P_{i,o}(\Delta \mathbf{s} | \mathbf{s}) - \sum_{\Delta \mathbf{s}: k \leq \ln n} f(a, b, r_i, r_o, k) q_{i,o}(\Delta \mathbf{s} | \mathbf{s}) \right| \\ & \lesssim \sum_{\Delta \mathbf{s}: k \leq \ln n} (\ln n)^2 |q_{i,o}(\Delta \mathbf{s} | \mathbf{s}) - P_{i,o}(\Delta \mathbf{s} | \mathbf{s})| \\ & \lesssim \frac{(\ln n)^8}{n} \sum_{\Delta \mathbf{s}} P_{i,o}(\Delta \mathbf{s} | \mathbf{s}) \leq \frac{(\ln n)^8}{n}. \end{aligned}$$

□

Importantly, for each $E_{\mathbf{s}}^q[f]$ we can find an explicit $\mathcal{E}_{\mathbf{s}}[f]$ such that

$$E_{\mathbf{s}}^q[f] = \mathcal{E}_{\mathbf{s}}[f] + O(1/n),$$

uniformly for $\mathbf{s} \in \mathbf{S}_\epsilon$. Here is how. Introduce the trivariate (probability) generating function of the parameters b, r_i, k with respect to q_i ,

$$F(x, y, w) := \sum_{b \geq 0} \sum_{r_i \geq 0} \sum_{k \geq \max\{1, b+r_i\}} x^b y^{r_i} w^k q_i(\Delta \mathbf{s} | \mathbf{s}).$$

Even though the definition of q_i in (3.13) is a bit forbidding, a simple algebraic computation—we encourage the reader to do it—shows that

$$\begin{aligned} F(x, y, w) &= \frac{\nu_i}{\nu_i + \nu_o} \frac{1}{e^{z_o} - 1} \left\{ \left(1 + \frac{z_i z_o x w}{\mu(e^{z_i} - 1)} \right)^{\nu_o} \right. \\ & \quad \times \exp \left(\frac{(\nu - \nu_i - \nu_o) z_i z_o y w}{(e^{z_i} - 1) \mu} + \frac{(\nu - \nu_i) z_i z_o w}{\mu} \right) - 1 \left. \right\}. \end{aligned}$$

Armed with this formula, we obviously can compute the q_i -expected values of the functions f in question by evaluating F and its partial derivatives at $x = y = w = 1$. To simplify the resulting q_i -expectations, we will use

$$\left(1 + \frac{z_i z_o}{\mu(e^{z_i} - 1)} \right)^{\nu_o} = \exp \left(\frac{\nu_o z_i z_o}{\mu(e^{z_i} - 1)} \right) + O(n^{-1}) \quad (3.15)$$

uniformly over $\mathbf{s} \in \mathbf{S}_\epsilon$. For ease of notation, each of the following sums are over $\Delta \mathbf{s}$ where $a = 1, r_o = 0, b \geq 0, r_i \geq 0$ and $k \geq \max\{1, b + r_i\}$. Necessarily, $q_i = 0$ unless $r_o = 0$; so $\sum_{\Delta \mathbf{s}} r_o q_i(\Delta \mathbf{s}|\mathbf{s}) = 0$. Further, $q_i = 0$ unless $a = 1$, in which case

$$\begin{aligned} \sum_{\Delta \mathbf{s}} a q_i(\Delta \mathbf{s}|\mathbf{s}) &= \sum_{\Delta \mathbf{s}} q_i(\Delta \mathbf{s}|\mathbf{s}) = F(1, 1, 1) \\ &= \frac{\nu_i}{\nu_i + \nu_o} \frac{1}{e^{z_o} - 1} \left\{ \exp \left(\frac{\nu_o z_i z_o}{(e^{z_i} - 1)\mu} + \frac{(\nu - \nu_i - \nu_o) z_i z_o}{(e^{z_i} - 1)\mu} \right. \right. \\ &\quad \left. \left. + \frac{(\nu - \nu_i) z_i z_o}{\mu} \right) - 1 \right\} + O(n^{-1}), \end{aligned}$$

uniformly over $\mathbf{s} \in \mathbf{S}_\epsilon$. We note that

$$\frac{\nu_o z_i z_o}{(e^{z_i} - 1)\mu} + \frac{(\nu - \nu_i - \nu_o) z_i z_o}{(e^{z_i} - 1)\mu} + \frac{(\nu - \nu_i) z_i z_o}{\mu} = \frac{(\nu - \nu_i) z_i z_o}{\mu} \left(\frac{1}{e^{z_i} - 1} + 1 \right) = z_o,$$

since $z_i e^{z_i} / (e^{z_i} - 1) = \ell(z_i) = \mu / (\nu - \nu_i)$. Hence

$$\sum_{\Delta \mathbf{s}} a q_i(\Delta \mathbf{s}|\mathbf{s}) = \frac{\nu_i}{\nu_i + \nu_o} + O(n^{-1}).$$

To estimate the q_i -averages of the remaining b, r_i, k , we evaluate the corresponding partial derivatives of F at $(1, 1, 1)$. For example,

$$\begin{aligned} \sum_{\Delta \mathbf{s}} b q_i(\Delta \mathbf{s}|\mathbf{s}) &= \sum_{\Delta \mathbf{s}} x^b y^{r_i} w^k q_i(\Delta \mathbf{s}|\mathbf{s}) \Big|_{x=y=w=1} = F_x(1, 1, 1) \\ &= \frac{\nu_i}{\nu_i + \nu_o} \frac{1}{e^{z_o} - 1} \left\{ \nu_o \left(1 + \frac{z_i z_o}{(e^{z_i} - 1)\mu} \right)^{\nu_o - 1} \frac{z_i z_o}{(e^{z_i} - 1)\mu} \right. \\ &\quad \left. \times \exp \left(\frac{(\nu - \nu_i - \nu_o) z_i z_o}{(e^{z_i} - 1)\mu} + \frac{(\nu - \nu_i) z_i z_o}{\mu} \right) \right\}. \end{aligned}$$

Using the asymptotic expression (3.15) as well as the definition of z_i and z_o , we can rewrite the above RHS as

$$\sum_{\Delta \mathbf{s}} b q_i(\Delta \mathbf{s}|\mathbf{s}) = \frac{\nu_i \nu_o \mu e^{-z_i}}{(\nu_i + \nu_o)(\nu - \nu_i)(\nu - \nu_o)} + O(n^{-1}).$$

Similarly

$$\sum_{\Delta \mathbf{s}} r_i q_i(\Delta \mathbf{s}|\mathbf{s}) = F_y(1, 1, 1) = \frac{\nu_i \mu (\nu - \nu_i - \nu_o) e^{-z_i}}{(\nu_i + \nu_o)(\nu - \nu_o)(\nu - \nu_i)} + O(n^{-1}),$$

and

$$\sum_{\Delta \mathbf{s}} k q_i(\Delta \mathbf{s}|\mathbf{s}) = F_w(1, 1, 1) = \frac{\nu_i \mu}{(\nu_i + \nu_o)(\nu - \nu_o)} + O(n^{-1}).$$

By symmetry between “in” and “out”, the q_o -average of a (resp. b) is found by switching i and o in the formula for the q_i -average of b (resp. a). Similarly, we determine the sums over q_o of the other variables. Thus, the $O(n^{-1})$ error terms aside, the leading terms for q -expectations of the transition parameters

$$\begin{aligned}
E_s^q[a] &= \mathcal{E}_s[a] + O(1/n); \quad \mathcal{E}_s[a] := \frac{\nu_i}{\nu_i + \nu_o} + \frac{\nu_i \nu_o \mu e^{-z_o}}{(\nu_i + \nu_o)(\nu - \nu_i)(\nu - \nu_o)}, \\
E_s^q[b] &= \mathcal{E}_s[b] + O(1/n); \quad \mathcal{E}_s[b] := \frac{\nu_o}{\nu_i + \nu_o} + \frac{\nu_i \nu_o \mu e^{-z_i}}{(\nu_i + \nu_o)(\nu - \nu_i)(\nu - \nu_o)}, \\
E_s^q[r_i] &= \mathcal{E}_s[r_i] + O(1/n); \quad \mathcal{E}_s[r_i] := \frac{\nu_i \mu (\nu - \nu_i - \nu_o) e^{-z_i}}{(\nu_i + \nu_o)(\nu - \nu_i)(\nu - \nu_o)}, \\
E_s^q[r_o] &= \mathcal{E}_s[r_o] + O(1/n); \quad \mathcal{E}_s[r_o] := \frac{\nu_o \mu (\nu - \nu_i - \nu_o) e^{-z_o}}{(\nu_i + \nu_o)(\nu - \nu_i)(\nu - \nu_o)}, \\
E_s^q[k] &= \mathcal{E}_s[k] + O(1/n); \quad \mathcal{E}_s[k] := \frac{\mu}{\nu_i + \nu_o} \left(\frac{\nu_o}{\nu - \nu_i} + \frac{\nu_i}{\nu - \nu_o} \right),
\end{aligned} \tag{3.16}$$

uniformly for $\mathbf{s} \in \mathbf{S}_\epsilon$ (\mathbf{S}_ϵ defined in (3.9)). Tellingly and importantly, all these functions are zero-degree homogeneous functions of \mathbf{s} . Combining Lemma 3.10 and (3.16), we have that uniformly over $\mathbf{s} \in \mathbf{S}_\epsilon$ and any $f = a, b, r_i, r_o, k$,

$$E_s[f] = \mathcal{E}_s[f] + O((\ln n)^8/n). \tag{3.17}$$

We naturally extend the approximate expectation to linear combinations of the terms a, b, r_i, r_o, k . For instance, $\mathcal{E}_s[\lambda_1 a + \lambda_2 b] := \lambda_1 \mathcal{E}_s[a] + \lambda_2 \mathcal{E}_s[b]$ for any constants λ_1 and λ_2 . Consequently, (3.17) holds for f being a linear combination of a, b, r_i, r_o, k .

The same technique works for the expected values of the 15 pairwise products a^2, ab, \dots, k^2 , via evaluating the second order partial derivatives of F at $(1, 1, 1)$. For instance, we find that q_i -averages of $f = a^2, b^2$ are given by $E_{s,i}[f] = \mathcal{E}_{s,i}[f] + O(1/n)$, where

$$\begin{aligned}
\mathcal{E}_{s,i}[a^2] &= \frac{\nu_i}{\nu_i + \nu_o}, \\
\mathcal{E}_{s,i}[b^2] &= \frac{\nu_i}{\nu_i + \nu_o} \frac{e^{z_o}}{e^{z_o} - 1} \left(\left(\frac{\nu_o}{e^{z_i} - 1} \frac{z_i z_o}{\mu} \right)^2 + \frac{\nu_o}{e^{z_i} - 1} \frac{z_i z_o}{\mu} \right).
\end{aligned}$$

Using symmetry, $\mathcal{E}_{s,o}[a^2]$ ($\mathcal{E}_{s,o}[b^2]$ resp.) is obtained from $\mathcal{E}_{s,i}[b^2]$ ($\mathcal{E}_{s,i}[a^2]$ resp.) by switching each instance of i with o , and o with i . Hence

$$\begin{aligned}
\mathcal{E}_s[a^2] &:= \mathcal{E}_{s,i}[a^2] + \mathcal{E}_{s,o}[a^2] \\
&= \frac{\nu_i}{\nu_i + \nu_o} + \frac{\nu_o}{\nu_i + \nu_o} \frac{e^{z_i}}{e^{z_i} - 1} \left(\left(\frac{\nu_i}{e^{z_o} - 1} \frac{z_i z_o}{\mu} \right)^2 + \frac{\nu_i}{e^{z_o} - 1} \frac{z_i z_o}{\mu} \right),
\end{aligned}$$

with $\mathcal{E}_s[b^2]$ obtained from the above expression by switching each i to o and each o to i . See Appendix A.1 for the full list of 15 approximate expected values. Further, each of

these approximate expectations are zero-degree homogeneous functions of \mathbf{s} . Needless to say, the asymptotic approximation (3.17) continues to hold, uniformly for $\mathbf{s} \in \mathbf{S}_\epsilon$, for all these f as well. Namely, we have that

Lemma 3.11. *Uniformly over $\mathbf{s} \in \mathbf{S}_\epsilon$ and any function f that is a linear combination of the linear terms a, b, r_i, r_o, k or their pairwise products a^2, ab, \dots, k^2 ,*

$$E_{\mathbf{s}}[f] = \mathcal{E}_{\mathbf{s}}[f] + O((\ln n)^8/n).$$

Here is a brief summary of what we have done so far. First, we described the Markovian deletion process which terminates at the (1,1)-core. Next we showed that we can lump together the digraph states with the same foursome $\{\mathbf{s}(t)\}$ without sacrificing the Markovian nature of the process. Then, we found a tractable asymptotic approximation $q(\Delta\mathbf{s}|\mathbf{s})$ of the transition probabilities $P(\Delta\mathbf{s}|\mathbf{s})$ for the states \mathbf{s} in the likely range of the random $\mathbf{s}(0)$ in $D(n, m = c_n n)$. Lastly we found sharp explicit approximations for the conditional expectations of the parameters a, b, r_i, r_o, k associated with the random transition from \mathbf{s} to the next state \mathbf{s}' , under condition $\mathbf{s} \in \mathbf{S}_\epsilon$. We will use these approximations shortly.

4 Characteristic Function of the Core Parameters

The deletion process $\{D(t)\}$ stops at $\bar{\tau} := \min\{t : \nu_i(t) = \nu_o(t) = 0\}$, the first time there are no semi-isolated vertices, i.e. $D(t) \equiv D(\bar{\tau})$ for $t \geq \bar{\tau}$. In terms of the reduced process $\{\mathbf{s}(t)\}$, $\mathbf{s}(t) \equiv \mathbf{s}(\bar{\tau})$ for $t \geq \bar{\tau}$. Let $\bar{\nu}$ and $\bar{\mu}$ denote the terminal number of vertices and arcs, i.e. $\bar{\nu} = \nu(\bar{\tau}), \bar{\mu} = \mu(\bar{\tau})$. Our ultimate goal is to show asymptotic Gaussianity of the pair $(\bar{\nu}, \bar{\mu})$ for the reduced process that starts at the random $\mathbf{s}(0)$ in $D(n, m = c_n n)$.

To this end, we show first that the pair $(\bar{\nu}, \bar{\mu})$ is asymptotically Gaussian for a deterministic $\mathbf{s}(0) = \mathbf{s}$ belonging to a likely range of initial states $\mathbf{s}(0)$ in $D(n, m = c_n n)$. Let $\varphi_{\mathbf{s}}(\mathbf{u})$, $\mathbf{u} = (u_1, u_2)^T \in \mathbb{R}^2$, denote the joint characteristic function of $(\bar{\nu}, \bar{\mu})$ for a generic, deterministic, $\mathbf{s}(0) = \mathbf{s}$. Formally, $\varphi_{\mathbf{s}}(\mathbf{u}) := E_{\mathbf{s}}[e^{iu_1\bar{\nu} + iu_2\bar{\mu}}]$, where as before, $E_{\mathbf{s}}[\cdot]$ denotes the expectation according to probability distribution of the process, $\{\mathbf{s}(t)\}$, starting at $\mathbf{s}(0) = \mathbf{s}$. Since the process is time-homogeneous, $\varphi_{\mathbf{s}}$ satisfies

$$\varphi_{\mathbf{s}}(\mathbf{u}) = \sum_{\mathbf{s}'} \varphi_{\mathbf{s}'}(\mathbf{u}) P(\mathbf{s}'|\mathbf{s}), \quad (4.1)$$

with $P(\mathbf{s}'|\mathbf{s})$ denoting the probability of one step-transition from \mathbf{s} to \mathbf{s}' . We would need to show existence of (smooth) functions $f_j, \psi_{j,k}$ ($j, k = 1, 2$), of the scaled \mathbf{s}/n , such that the 2×2 symmetric matrix $\boldsymbol{\psi} = \{\psi_{j,k}\}_{1 \leq j,k \leq 2}$ is positive-definite, and

$$G_n(\mathbf{s}/n, \mathbf{u}) := \exp\left(in\mathbf{u}^T \mathbf{f}(\mathbf{s}/n) - \frac{n}{2}\mathbf{u}^T \boldsymbol{\psi}(\mathbf{s}/n)\mathbf{u}\right), \quad \mathbf{f}^T := (f_1, f_2), \quad (4.2)$$

sharply approximates $\varphi_{\mathbf{s}}(\mathbf{u})$ for $\|\mathbf{u}\| = O(n^{-1/2})$ (this order of \mathbf{u} is dictated by our goal of proving asymptotic normality of the numbers of vertices and arcs in the (1,1)-core

centered by their respective means and scaled by $n^{1/2}$, the asymptotic order of their standard deviations). Necessarily, such f_j and $\psi_{j,k}$ would have to satisfy the boundary conditions

$$f_1(\alpha, 0, 0, \gamma) = \alpha, \quad f_2(\alpha, 0, 0, \gamma) = \gamma, \quad \psi_{j,k}(\alpha, 0, 0, \gamma) = 0. \quad (4.3)$$

Indeed, if ν_i and ν_o are both zero, our process is already stopped at time $t = 0$, and so $\bar{\nu} = \nu$ and $\bar{\mu} = \mu$. Ideally, we would hope to determine \mathbf{f} and $\boldsymbol{\psi}$ out of the condition that $G_n(\mathbf{s}/n, \mathbf{u})$ “almost” satisfies the equation (4.1), with “almost” to be specified shortly.

This method of proving asymptotic normality with simultaneous determination of the attendant parameters had been used earlier, Pittel [31], [29], and Pittel and Weishaar [34].

However, our analysis would inevitably rely on the asymptotic approximation of the transition probability $P(\Delta \mathbf{s} | \mathbf{s})$ by the sub-stochastic $q(\Delta \mathbf{s} | \mathbf{s})$, established *only* for $\mathbf{s} \in \mathbf{S}_\epsilon$, defined in (3.9); see Lemma 3.8, Lemma 3.9. So effectively we are forced to stop the process possibly earlier, at time

$$\hat{\tau} = \begin{cases} \min\{t < \bar{\tau} : \mathbf{s}(t) \notin \mathbf{S}_\epsilon\}, \\ \bar{\tau}, & \text{if no such } t \text{ exists.} \end{cases} \quad (4.4)$$

and to consider instead $(\hat{\nu}, \hat{\mu}) := (\nu(\hat{\tau}), \mu(\hat{\tau}))$. We will show, however, that, for $\mathbf{s}(0)$ from a large enough subset of \mathbf{S}_ϵ , a.a.s. $\hat{\tau} = \bar{\tau}$. So asymptotic normality of $(\hat{\nu}, \hat{\mu})$ will imply asymptotic normality of $(\bar{\nu}, \bar{\mu})$.

We will use \hat{P} and $\hat{E}_\mathbf{s}[\cdot]$ for the transition probabilities and the expectations, starting at generic state \mathbf{s} , for this (still Markovian) modification of the deletion process that freezes at time $\hat{\tau}$. The corresponding $\hat{\varphi}_\mathbf{s}(\mathbf{u}) := \hat{E}_\mathbf{s}[e^{iu_1\hat{\nu} + iu_2\hat{\mu}}]$ satisfies

$$\hat{\varphi}_\mathbf{s}(\mathbf{u}) = \sum_{\mathbf{s}'} \hat{\varphi}_{\mathbf{s}'}(\mathbf{u}) \hat{P}(\mathbf{s}(1) = \mathbf{s}' | \mathbf{s}(0) = \mathbf{s}). \quad (4.5)$$

Thus we set up to determine smooth functions f_j and $\psi_{j,k}$, for $1 \leq j, k \leq 2$, such that for $\|\mathbf{u}\| = O(n^{-1/2})$, and $\mathbf{s} \in \mathbf{S}_\epsilon$, the function $G_n(\mathbf{s}/n, \mathbf{u})$ defined in (4.2) almost satisfies (4.5), namely

$$G_n(\mathbf{s}/n, \mathbf{u}) - \sum_{\mathbf{s}'} G_n(\mathbf{s}'/n, \mathbf{u}) \hat{P}(\mathbf{s}(1) = \mathbf{s}' | \mathbf{s}(0) = \mathbf{s}) = o(1/n). \quad (4.6)$$

We need this additive error term to be that small as at one point we will have to sum up those terms over the duration of the process.

To this end, we plug $G_n(\cdot, \mathbf{u})$ into the sum in (4.6) and simplify the resulting expression via Taylor-expanding the generic exponents around \mathbf{s} . That’s where we need and hope for sufficient smoothness—two continuous derivatives would do—of the functions \mathbf{f} and $\boldsymbol{\psi}$. Requiring the resulting LHS of (4.6) be $o(1/n)$, we will arrive at a system of first-order PDF for \mathbf{f} and $\boldsymbol{\psi}$ and solve it using the ODE for the characteristics of the PDE, thus establishing existence of the required functions.

Here are the details. Recall that \mathbf{s}' can follow from \mathbf{s} only if

$$\Delta \mathbf{s} = \mathbf{s}' - \mathbf{s} = (-a - b, r_i - a, r_o - b, -k)$$

is such that $k \geq \max\{a + r_o, b + r_i\}$ and either $a = 1, r_o = 0$ or $b = 1, r_i = 0$. Suppose that $k \leq \ln n$. Assume that the second order derivatives of \mathbf{f} and $\boldsymbol{\psi}$ have absolute values bounded on the line segment $[\mathbf{s}, \mathbf{s}']$. Then, denoting the gradient of f_j at \mathbf{s} by ∇f_j , we have

$$\begin{aligned} nu_j f_j(\mathbf{s}'/n) &= nu_j f_j(\mathbf{s}/n) + nu_j (\Delta \mathbf{s}/n)^T \nabla f_j + O(n|u_j|(\ln n)^2 n^{-2}) \\ &= nu_j f_j(\mathbf{s}/n) + u_j \Delta \mathbf{s}^T \nabla f_j + O((\ln n)^2 n^{-3/2}) \\ &= nu_j f_j(\mathbf{s}/n) + O((\ln n)n^{-1/2}), \end{aligned}$$

as $\|\mathbf{u}\| = O(n^{-1/2})$. Likewise

$$\begin{aligned} nu_j u_k \psi_{j,k}(\mathbf{s}'/n) &= nu_j u_k \psi_{j,k}(\mathbf{s}/n) + u_j u_k \Delta \mathbf{s}^T \nabla \psi_{j,k} + O((\ln n)^2 n^{-2}) \\ &= nu_j u_k \psi_{j,k}(\mathbf{s}/n) + O((\ln n)n^{-1}). \end{aligned}$$

Using multiplicativity of e^z , and $e^y = 1 + y + y^2/2 + O(|y|^3)$, we have then: for $\|\mathbf{u}\| = O(n^{-1/2})$,

$$\begin{aligned} \frac{G_n(\mathbf{s}'/n, \mathbf{u})}{G_n(\mathbf{s}/n, \mathbf{u})} &= \exp \left(i \sum_j u_j \Delta \mathbf{s}^T \nabla f_j - \frac{1}{2} \sum_{j,k} u_j u_k \Delta \mathbf{s}^T \nabla \psi_{j,k} + O \left(\frac{(\ln n)^2}{n^{3/2}} \right) \right) \\ &= 1 + i \sum_j u_j \Delta \mathbf{s}^T \nabla f_j - \frac{1}{2} \sum_{j,k} u_j u_k [(\Delta \mathbf{s}^T \nabla f_j)(\Delta \mathbf{s}^T \nabla f_k) + \Delta \mathbf{s}^T \nabla \psi_{j,k}] + O \left(\frac{(\ln n)^2}{n^{3/2}} \right). \end{aligned}$$

Here the second and the third term are of order $n^{-1/2}$ and n^{-1} respectively. So, for G_n to satisfy (4.6) for all \mathbf{u} in question, it would suffice that the sums of the second and third terms over $\Delta \mathbf{s}$ are $o(n^{-1/2})$ and $o(1)$, respectively. In other words,

$$\left(\sum_{\Delta \mathbf{s}} \Delta \mathbf{s}^T \hat{P}(\Delta \mathbf{s}|\mathbf{s}) \right) \nabla f_j = \hat{E}_{\mathbf{s}}[\Delta \mathbf{s}^T] \nabla f_j = o(n^{-1/2}), \quad (4.7)$$

and

$$\hat{E}_{\mathbf{s}}[\Delta \mathbf{s}^T] \nabla \psi_{j,k} + \hat{E}_{\mathbf{s}}[(\Delta \mathbf{s}^T \nabla f_j)(\Delta \mathbf{s}^T \nabla f_k)] = o(1). \quad (4.8)$$

The reason we haven't replaced those $o(n^{-1/2})$ by 0 is that we still have to approximate the expected values in these two equations using Lemma 3.10. At any rate, these are PDE-type equations for \mathbf{f} and $\boldsymbol{\psi}$ that have to be solved under the boundary conditions (4.3).

5 PDE

5.1 PDE for the mean parameters

For the PDE-type equation (4.7), we need to estimate

$$\hat{E}_{\mathbf{s}}[\Delta \mathbf{s}^T] = (\hat{E}_{\mathbf{s}}[-a - b], \hat{E}_{\mathbf{s}}[r_i - a], \hat{E}_{\mathbf{s}}[r_o - b], \hat{E}_{\mathbf{s}}[-k]).$$

By the definition of \hat{P} , $\hat{E}_s[\Delta \mathbf{s}^T] = \mathbf{0}$, if $\mathbf{s} \notin \mathbf{S}_\epsilon$, and $\hat{E}_s[\Delta \mathbf{s}^T] = E_s[\Delta \mathbf{s}^T]$ for $\mathbf{s} \in \mathbf{S}_\epsilon$. By Lemma 3.11, we have

$$E_s[\Delta \mathbf{s}^T] = \mathcal{E}_s[\Delta \mathbf{s}^T] + O((\ln n)^8/n), \quad (5.1)$$

uniformly for $\mathbf{s} \in \mathbf{S}_\epsilon$. The four coefficients by the partial derivatives of f_j are computed, by linearity, via (3.16).

In order to establish the existence of f_j satisfying (4.7), it suffices to prove existence of f_j that satisfy the *homogeneous* partial differential equation:

PDE:

$$\mathcal{E}_s[-(a+b)](f_j)_\alpha + \mathcal{E}_s[r_i - a](f_j)_{\beta_i} + \mathcal{E}_s[r_o - b](f_j)_{\beta_o} + \mathcal{E}_s[-k](f_j)_\gamma = 0, \quad (5.2)$$

$\mathbf{s} := n(\alpha, \beta_i, \beta_o, \gamma)$, under the boundary conditions: $f_1(\alpha, 0, 0, \gamma) = \alpha$ and $f_2(\alpha, 0, 0, \gamma) = \gamma$. This PDE equation is zero-degree homogeneous, since the coefficients depend on $(\alpha, \beta_i, \beta_o, \gamma)$ only.

5.2 The Characteristics for the PDE

The characteristics of the first-order PDE (5.2) are the trajectories of the following system of ODEs:

$$\frac{d\alpha}{dt} = \mathcal{E}_s[-a-b], \quad \frac{d\beta_i}{dt} = \mathcal{E}_s[r_i - a], \quad \frac{d\beta_o}{dt} = \mathcal{E}_s[r_o - b], \quad \frac{d\gamma}{dt} = \mathcal{E}_s[-k].$$

Using (3.16), after minor simplifications we obtain the ODE in an explicit form:

$$\begin{aligned} \frac{d\alpha}{dt} &= -1 - \frac{\beta_i \beta_o \gamma (e^{-z_i} + e^{-z_o})}{(\beta_i + \beta_o)(\alpha - \beta_i)(\alpha - \beta_o)}, \\ \frac{d\beta_i}{dt} &= \frac{\beta_i}{\beta_i + \beta_o} \left(\frac{\gamma(\alpha - \beta_i - \beta_o) e^{-z_i}}{(\alpha - \beta_i)(\alpha - \beta_o)} - 1 - \frac{\beta_o \gamma e^{-z_o}}{(\alpha - \beta_i)(\alpha - \beta_o)} \right), \\ \frac{d\beta_o}{dt} &= \frac{\beta_o}{\beta_i + \beta_o} \left(\frac{\gamma(\alpha - \beta_i - \beta_o) e^{-z_o}}{(\alpha - \beta_i)(\alpha - \beta_o)} - 1 - \frac{\beta_i \gamma e^{-z_i}}{(\alpha - \beta_i)(\alpha - \beta_o)} \right), \\ \frac{d\gamma}{dt} &= -\frac{\gamma}{\beta_i + \beta_o} \left(\frac{\beta_o}{\alpha - \beta_i} + \frac{\beta_i}{\alpha - \beta_o} \right). \end{aligned} \quad (5.3)$$

Here $z_i = z_i(\alpha, \beta_i, \beta_o, \gamma)$ and $z_o = z_o(\alpha, \beta_i, \beta_o, \gamma)$ are defined by

$$\ell(z_i) := \frac{z_i e^{z_i}}{e^{z_i} - 1} = \frac{\gamma}{\alpha - \beta_i}, \quad \ell(z_o) := \frac{z_o e^{z_o}}{e^{z_o} - 1} = \frac{\gamma}{\alpha - \beta_o}. \quad (5.4)$$

Note that this fits with our previous definition of $z_i(\mathbf{s})$ and $z_o(\mathbf{s})$ (Lemma 3.3) since the RHS of (5.4) is zero-degree homogeneous. The differential equations above are certainly well defined for $\mathbf{w} = (\alpha, \beta_i, \beta_o, \gamma) \geq \mathbf{0}$ satisfying

$$\beta_i, \beta_o > 0; \beta_i + \beta_o < \alpha; \quad \frac{\gamma}{\alpha - \beta_i}, \frac{\gamma}{\alpha - \beta_o} > 1. \quad (5.5)$$

This seemingly formidable system has two remarkable integrals, the explicit state-functions that remain constant along every characteristic.

In vector notation, this system becomes

$$\frac{d\mathbf{w}}{dt} = \mathcal{E}_s[\Delta \mathbf{s}] =: \mathbf{H}(\mathbf{w}). \quad (5.6)$$

Suppose \mathbf{w}_0 is such that $\mathbf{H}(\mathbf{w})$ has bounded partial derivatives in a ball centered at \mathbf{w}_0 . Then the system (5.6) has a unique solution $\mathbf{w}(t)$, $\mathbf{w}(0) = \mathbf{w}_0$, defined for $t \in [0, t(\mathbf{w}_0)]$, for some $t(\mathbf{w}_0) > 0$ ($t(\mathbf{w}_0)$ could be infinite).

Proposition 5.1. *Let $I_2(\mathbf{w}) = z_i z_o / \gamma$. For $t \leq t(\mathbf{w}_0)$, we have $I_2(\mathbf{w}(t)) \equiv I_2(\mathbf{w}(0))$.*

Proof. We need to compute dz_i/dt and dz_o/dt . By (5.4),

$$\frac{dz_i}{dt} = \frac{1}{k'(z_i)} \cdot \frac{d}{dt} \left(\frac{\gamma}{\alpha - \beta_i} \right) = \frac{1}{k'(z_i)} \cdot \frac{\gamma'(\alpha - \beta_i) - \gamma(\alpha' - \beta'_i)}{(\alpha - \beta_i)^2}.$$

Plugging in the expressions for α' , β'_i and γ' from (5.3) and simplifying, we arrive at a surprisingly simple formula,

$$\frac{dz_i}{dt} = -\frac{\beta_i z_i}{(\beta_i + \beta_o)(\alpha - \beta_o)}, \quad (5.7)$$

Likewise

$$\frac{dz_o}{dt} = -\frac{\beta_o z_o}{(\beta_i + \beta_o)(\alpha - \beta_i)}. \quad (5.8)$$

Using the formula for $d\gamma/dt$ from (5.3), we see that

$$\frac{1}{z_i} \frac{dz_i}{dt} + \frac{1}{z_o} \frac{dz_o}{dt} - \frac{1}{\gamma} \frac{d\gamma}{dt} = 0,$$

or simply

$$\frac{d}{dt} [\ln z_i + \ln z_o - \ln \gamma] = 0 \implies \frac{z_i(t) z_o(t)}{\gamma(t)} \equiv \text{constant}.$$

□

Proposition 5.2. *Let $I_1(\mathbf{w}) = \gamma(\alpha - \beta_i - \beta_o)/[(\alpha - \beta_i)(\alpha - \beta_o)]$. For $t \leq t(\mathbf{w}_0)$, we have $I_1(\mathbf{w}(t)) \equiv I_1(\mathbf{w}(0))$.*

Proof. A straightforward computation of $d(\alpha - \beta_i - \beta_o)/dt$ from (5.3), followed by minor simplification, yields

$$\frac{1}{\alpha - \beta_i - \beta_o} \frac{d(\alpha - \beta_i - \beta_o)}{dt} = \frac{1}{e^{z_i} - 1} \frac{dz_i}{dt} + \frac{1}{e^{z_o} - 1} \frac{dz_o}{dt},$$

which can be rewritten as

$$\frac{d}{dt} [\ln(\alpha - \beta_i - \beta_o) - \ln(1 - e^{-z_i}) - \ln(1 - e^{-z_o})] = 0.$$

So

$$\ln \left[(\alpha - \beta_i - \beta_o) \frac{e^{z_i}}{e^{z_i} - 1} \frac{e^{z_o}}{e^{z_o} - 1} \right] \equiv \text{constant}.$$

Using the definition of z_i and z_o , we see that

$$(\alpha - \beta_i - \beta_o) \frac{e^{z_i}}{e^{z_i} - 1} \frac{e^{z_o}}{e^{z_o} - 1} = \frac{\alpha - \beta_i - \beta_o}{z_i z_o} \frac{\gamma^2}{(\alpha - \beta_i)(\alpha - \beta_o)} = \frac{\gamma}{z_i z_o} \frac{\gamma(\alpha - \beta_i - \beta_o)}{(\alpha - \beta_i)(\alpha - \beta_o)}.$$

By the previous proposition, $\gamma/(z_i z_o)$ is constant, so $\gamma(\alpha - \beta_i - \beta_o)/[(\alpha - \beta_i)(\alpha - \beta_o)]$ is constant as well. \square

Thus we have proved that $I_1(\mathbf{w}(t))$ and $I_2(\mathbf{w}(t))$ are constant on the trajectory starting at \mathbf{w}_0 . As customary, we call these functions the integrals of the ODE (5.6). The constancy of I_j along the trajectory means that the gradient of I_j is orthogonal to the “velocity”, $H(\mathbf{w})$, along the trajectory, (5.3), i.e.

$$\nabla I_j(\mathbf{w}) \cdot \mathcal{E}_s[\Delta \mathbf{s}] = 0, \quad (5.9)$$

where $\mathbf{s}/n = \mathbf{w}$.

Observe that $I_1(\mathbf{w})$, $I_2(\mathbf{w})$ are $F_1(\mathbf{s})$, $F_2(\mathbf{s})$ (defined in (3.7)), with $\mathbf{w} = \mathbf{s}/n$. The perfect constancy of $I_j(\mathbf{w}(t))$ along the characteristics of the PDE predicts that $F_j(\mathbf{s}(t)/n)$ will be proven to be, a.a.s., almost constant along the random realization of the deletion process. That explains why we used $F_j(\mathbf{s})$ in the definition of \mathbf{S}_ϵ , the set of good \mathbf{s} .

Continuing with the characteristics, we will prove existence *at large* of a solution $\mathbf{w}(t)$ of (5.6) for a “good” starting point \mathbf{w}_0 . Analogously to good \mathbf{s} , i.e. $\mathbf{s} \in \mathbf{S}_\epsilon$, the (open) set of good \mathbf{w} is defined by

$$\mathbf{W}_\epsilon := \left\{ \mathbf{w} \geq \mathbf{0} : \mathbf{w} \text{ meets (5.5); } I_1(\mathbf{w}), I_2(\mathbf{w}) \in (c_n - \epsilon, c_n + \epsilon) \right\}; \quad (5.10)$$

so if $\mathbf{s} \in \mathbf{S}_\epsilon$, then $\mathbf{s}/n \in \mathbf{W}_\epsilon$.

Fact 5.3. *There exists a fixed $\delta_1 > 0$ such that uniformly for $\mathbf{w} \in \mathbf{W}_\epsilon$*

- (i) $\alpha, \alpha - \beta_i, \alpha - \beta_o, \alpha - \beta_i - \beta_o \geq \delta_1$;
- (ii) $\gamma - \alpha \geq \delta_1$;
- (iii) $z_i, z_o \in [\delta_1, 1/\delta_1]$.

We omit the proof since it differs only superficially from the proof of Fact 3.7.

Furthermore, for $\mathbf{w}_0 \in \mathbf{W}_\epsilon$, there exists a unique solution of (5.6), satisfying $\mathbf{w}(0) = \mathbf{w}_0$, defined for $t \in [0, \mathcal{T})$ where $\mathcal{T} = \mathcal{T}(\mathbf{w}_0) = \sup\{t : \mathbf{w}(t') \in \mathbf{W}_\epsilon \text{ for } t' \leq t\}$. Indeed, by Fact 5.3, on compact subsets of the interior of \mathbf{W}_ϵ , the partial derivatives of $H(\mathbf{w})$ in (5.3) are bounded; it remains to use the existence and uniqueness Theorem for systems of ODE's and a standard compactness argument.

Now, $\mathcal{T} = \infty$ means that the trajectory $\{\mathbf{w}(t)\}$ stays in \mathbf{W}_ϵ indefinitely. To determine f_1 and f_2 , the solution of the PDE (5.2), meeting the boundary conditions $f_1(\alpha, 0, 0, \gamma) =$

α , $f_2(\alpha, 0, 0, \gamma) = \gamma$, we need to show that the characteristic of the PDE, i.e. the solution of (5.6), reaches $\partial(\mathbf{W}_\epsilon)$, the boundary of \mathbf{W}_ϵ , in finite time $\mathcal{T} < \infty$ and that $\beta_i(\mathcal{T}) = \beta_o(\mathcal{T}) = 0$.

To motivate the proof that the trajectory ends on this desired range of the boundary, let's see what happens when we end there. Suppose that $\mathcal{T} < \infty$ and moreover that $\beta_i(\mathcal{T}) = \beta_o(\mathcal{T}) = 0$, and $\alpha(\mathcal{T}), \gamma(\mathcal{T}) > 0$. Then

$$I_1(\mathcal{T}) = \frac{\gamma(\mathcal{T})}{\alpha(\mathcal{T}) - \beta_i(\mathcal{T})} = \frac{\gamma(\mathcal{T})}{\alpha(\mathcal{T}) - \beta_o(\mathcal{T})} = \frac{\gamma(\mathcal{T})}{\alpha(\mathcal{T})}.$$

So in this case, $z_i(t)$ and $z_o(t)$ both end with the same values! In fact, $z_i(\mathcal{T}) = z_o(\mathcal{T}) = z(\gamma(\mathcal{T})/\alpha(\mathcal{T}))$ where, as before, $z(\eta)$ is the unique positive root of $\ell(z) = \eta$ (definition of $\ell(z)$ in (5.4)). Since $I_1(\mathcal{T}) = I_1(\mathbf{w}(0))$, we see then that

$$z_i(\mathcal{T}) = z_o(\mathcal{T}) = z(\eta(\mathbf{w}_0)), \quad \eta(\mathbf{w}) := \frac{\gamma(\alpha - \beta_i - \beta_o)}{(\alpha - \beta_i)(\alpha - \beta_o)}. \quad (5.11)$$

In light of this preliminary discussion, we introduce $z^* = z^*(\mathbf{w}) = z(\eta(\mathbf{w}))$. We know that z^* is constant along the trajectory, so z^* depends only on \mathbf{w}_0 . Since

$$\frac{\gamma(\alpha - \beta_i - \beta_o)}{(\alpha - \beta_i)(\alpha - \beta_o)} \leq \min \left\{ \frac{\gamma}{\alpha - \beta_i}, \frac{\gamma}{\alpha - \beta_o} \right\},$$

and $z(\eta)$ is increasing, we have that $z_i, z_o \geq z^*$ along the trajectory. Note that z^* is bounded away from 0 and ∞ because $\eta(\mathbf{w}_0)$ is bounded away from 1 and ∞ uniformly for $\mathbf{w}_0 \in \mathbf{W}_\epsilon$.

Proposition 5.4. *If $\mathbf{w}_0 \in \mathbf{W}_\epsilon$, then $\beta_i(t)/\beta_o(t)$ is monotone, varying toward 1.*

Proof. First, from the second line of (5.3),

$$\frac{1}{\beta_i} \frac{d\beta_i}{dt} = \frac{\gamma(\alpha - \beta_i - \beta_o)e^{-z_i}}{(\alpha - \beta_i)(\alpha - \beta_o)} - \frac{1}{\beta_i + \beta_o} - \frac{\beta_o \gamma e^{-z_o}}{(\beta_i + \beta_o)(\alpha - \beta_i)(\alpha - \beta_o)}.$$

Using $\ell(z_i) = \gamma/(\alpha - \beta_i)$ and $\ell(z_o) = \gamma/(\alpha - \beta_o)$, we find that

$$\frac{1}{\beta_i} \frac{d\beta_i}{dt} = \frac{(\alpha - \beta_i - \beta_o)z_i}{(\beta_i + \beta_o)(\alpha - \beta_o)(e^{z_i} - 1)} - \frac{1}{\beta_i + \beta_o} - \frac{\beta_o z_o}{(\beta_i + \beta_o)(\alpha - \beta_i)(e^{z_o} - 1)}, \quad (5.12)$$

and similarly we obtain

$$\frac{1}{\beta_o} \frac{d\beta_o}{dt} = \frac{(\alpha - \beta_i - \beta_o)z_o}{(\beta_i + \beta_o)(\alpha - \beta_i)(e^{z_o} - 1)} - \frac{1}{\beta_i + \beta_o} - \frac{\beta_i z_i}{(\beta_i + \beta_o)(\alpha - \beta_o)(e^{z_i} - 1)}.$$

By subtracting these last two equations, we find that

$$\begin{aligned} \frac{1}{\beta_i} \frac{d\beta_i}{dt} - \frac{1}{\beta_o} \frac{d\beta_o}{dt} &= \frac{1}{(\beta_i + \beta_o)(\alpha - \beta_o)(e^{z_i} - 1)} \{(\alpha - \beta_i - \beta_o)z_i + \beta_i z_i\} \\ &\quad - \frac{1}{(\beta_i + \beta_o)(\alpha - \beta_i)(e^{z_o} - 1)} \{(\alpha - \beta_i - \beta_o)z_o + \beta_o z_o\}. \end{aligned}$$

Therefore

$$\frac{d}{dt} \ln \left(\frac{\beta_i}{\beta_o} \right) = \frac{1}{\beta_i + \beta_o} \left(\frac{z_i}{e^{z_i} - 1} - \frac{z_o}{e^{z_o} - 1} \right).$$

Since $\ell(z)$ is strictly increasing, and $\gamma/(\alpha - \beta)$ is increasing with β , we see that $\beta_i < \beta_o$ iff $z_i < z_o$. Besides, $x/(e^x - 1)$ decreases for $x > 0$. Therefore if $\beta_i/\beta_o < 1$, then $\frac{d}{dt} \ln \beta_i/\beta_o > 0$, and if $\beta_i/\beta_o > 1$, then $\frac{d}{dt} \ln \beta_i/\beta_o < 0$. \square

Proposition 5.5. *If $\mathbf{w}_0 \in \mathbf{W}_\epsilon$, then the trajectory $\mathbf{w}(t)$ starting at \mathbf{w}_0 satisfies (i) each of $\alpha, \beta_i, \beta_o, \gamma, z_i$ and z_o is decreasing with t , and (ii) $\mathcal{T} < \infty$.*

Proof. (i) That α, γ, z_i and z_o decrease with t is immediate from the ODE (5.3) as well as (5.7) and (5.8). Further, from (5.12) and $z_i \geq z^*$, where z^* is bounded away from zero, there exists a fixed $\delta > 0$, such that

$$\frac{1}{\beta_i} \frac{d\beta_i}{dt} \leq \frac{1}{\beta_i + \beta_o} \left(\frac{z^*}{e^{z^*} - 1} - 1 \right) \leq -\frac{\delta}{\beta_i + \beta_o}.$$

Therefore $\beta_i(t)$ is decreasing, and so is $\beta_o(t)$, since likewise

$$\frac{1}{\beta_i} \frac{d\beta_i}{dt} \leq -\frac{\delta}{\beta_i + \beta_o}.$$

(ii) From part (i) it follows that $d(\beta_i + \beta_o)/dt \leq -\delta$. Therefore

$$\mathcal{T} \leq (\beta_i(0) + \beta_o(0))/\delta < \infty.$$

\square

Thus \mathcal{T} is finite and $\mathbf{w}(\mathcal{T}-) := \lim_{t \rightarrow \mathcal{T}-} \mathbf{w}(t)$ exists. Defining $\mathbf{w}(\mathcal{T}) = \mathbf{w}(\mathcal{T}-)$, we make $\mathbf{w}(t)$ continuous on the closed interval $[0, \mathcal{T}]$.

Lemma 5.6. *Let $\mathbf{w}_0 \in \mathbf{W}_\epsilon$. Then $\mathbf{w}(\mathcal{T}) = (\alpha(\mathcal{T}), 0, 0, \gamma(\mathcal{T}))$, with $\alpha(\mathcal{T}) > 0$ and $\gamma(\mathcal{T}) > 0$.*

Proof. We know that each coordinate of $\mathbf{w}(t)$ is non-negative and decreasing. So, by the definition (5.10) of \mathbf{W}_ϵ , the condition $\mathbf{w}(\mathcal{T}) \in \partial(\mathbf{W}_\epsilon)$ means that at $t = \mathcal{T}$

- either at least one of α, β_i, β_o and γ is zero,
- or the condition (5.5) is violated.

Indeed, since $I_j(\mathbf{w}(t)) \equiv \text{const}$ for $t < \mathcal{T}$, $I_j(\mathbf{w}(\mathcal{T}))$ is still in $(c_n - \epsilon, c_n + \epsilon)$, $j = 1, 2$. Further, by Fact 5.3, $\alpha - \beta_i - \beta_o$ is bounded away from zero uniformly for $\mathbf{w} \in \mathbf{W}_\epsilon$. Thus the case $\alpha - \beta_i - \beta_o = 0$ at $t = \mathcal{T}$ is ruled out, and so $\alpha > \max\{\beta_i, \beta_o\}$. Further, as $z_i, z_o \in [z^*, \max\{z_i(0), z_o(0)\}]$, we see that $\ell(z_i) = \gamma/(\alpha - \beta_i)$ and $\ell(z_o) = \gamma/(\alpha - \beta_o)$ which are both at least $\ell(z^*) \in (c_n - \epsilon, c_n + \epsilon)$ must also be greater than 1. Thus the only reason that $\mathbf{w}(\mathcal{T}) \notin \mathbf{W}_\epsilon$ is that at least one of $\beta_i(\mathcal{T}), \beta_o(\mathcal{T})$ is zero. According to Proposition 5.4, $\beta_i(t)/\beta_o(t)$ varies monotonically in the direction towards 1, so we conclude that both $\beta_i(\mathcal{T})$ and $\beta_o(\mathcal{T})$ are zero. \square

5.3 Determining the mean parameters

The study of the characteristics of the PDE (5.2) in the previous section allows us to give explicit formulas for the mean parameters f_1 and f_2 .

Proposition 5.7. *For $\mathbf{w} = (\alpha, \beta_i, \beta_o, \gamma) \in \mathbf{W}_\epsilon$, the functions*

$$f_1(\alpha, \beta_i, \beta_o, \gamma) = \frac{z \left(\frac{\gamma(\alpha - \beta_i - \beta_o)}{(\alpha - \beta_i)(\alpha - \beta_o)} \right)^2}{z \left(\frac{\gamma}{\alpha - \beta_i} \right) z \left(\frac{\gamma}{\alpha - \beta_o} \right)} \cdot \frac{(\alpha - \beta_i)(\alpha - \beta_o)}{\alpha - \beta_i - \beta_o},$$

$$f_2(\alpha, \beta_i, \beta_o, \gamma) = \frac{z \left(\frac{\gamma(\alpha - \beta_i - \beta_o)}{(\alpha - \beta_i)(\alpha - \beta_o)} \right)^2}{z \left(\frac{\gamma}{\alpha - \beta_i} \right) z \left(\frac{\gamma}{\alpha - \beta_o} \right)} \cdot \gamma,$$

solves the PDE (5.2), where, we recall, $z(\eta)$ is the unique positive root of $\ell(z) = \eta$. Consequently, f_1 and f_2 are smooth on \mathbf{W}_ϵ , meaning that, on \mathbf{W}_ϵ , the functions f_1, f_2 have partial derivatives of every order, and for each fixed ℓ , the ℓ -order derivatives are uniformly bounded.

Proof. Let $\mathbf{w}(0) \in \mathbf{W}_\epsilon$. By Lemma 5.6, we have that $\mathbf{w}(\mathcal{T}) = (\alpha(\mathcal{T}), 0, 0, \gamma(\mathcal{T}))$, where $\alpha(\mathcal{T}), \gamma(\mathcal{T}) > 0$. Along the characteristic, the functions $f_j(\mathbf{w})$ each have a constant value since the PDE's (5.2) for f_1 and f_2 are homogeneous. By the boundary conditions for f_1 and f_2 , we must have then that

$$f_1(\mathbf{w}(0)) = f_1(\mathbf{w}(\mathcal{T})) = \alpha(\mathcal{T}), \quad f_2(\mathbf{w}(0)) = f_2(\mathbf{w}(\mathcal{T})) = \gamma(\mathcal{T}).$$

It remains to find $\alpha(\mathcal{T})$ and $\gamma(\mathcal{T})$ as functions of $\mathbf{w}(0)$.

Since $I_1(\mathbf{w}(t)) = \gamma(\alpha - \beta_i - \beta_o)/[(\alpha - \beta_i)(\alpha - \beta_o)] \equiv \text{const}$, we obtain

$$\frac{\gamma(\mathcal{T})}{\alpha(\mathcal{T})} = \frac{\gamma(\mathcal{T})(\alpha(\mathcal{T}) - 0 - 0)}{(\alpha(\mathcal{T}) - 0)(\alpha(\mathcal{T}) - 0)} = \frac{\gamma(0)(\alpha(0) - \beta_i(0) - \beta_o(0))}{(\alpha(0) - \beta_i(0))(\alpha(0) - \beta_o(0))}. \quad (5.13)$$

Furthermore, by (5.11), $z_i(\mathcal{T}) = z_o(\mathcal{T}) = z^*$, where

$$z^* = z(I_1(\mathbf{w}(\mathcal{T}))) = z(I_1(\mathbf{w}(0))),$$

since $I_1(\mathbf{w}(t)) \equiv I_1(\mathbf{w}(0))$. The constancy of $I_2(\mathbf{w}) = z_i z_o / \gamma$ along the trajectory implies that

$$\gamma(\mathcal{T}) = \gamma(0) \frac{z_i(\mathcal{T}) z_o(\mathcal{T})}{z_i(0) z_o(0)} = \gamma(0) \frac{(z^*)^2}{z_i(0) z_o(0)}. \quad (5.14)$$

It follows from (5.13) and (5.14) that

$$\alpha(\mathcal{T}) = \frac{(z^*)^2}{z_i(0) z_o(0)} \frac{(\alpha(0) - \beta_i(0))(\alpha(0) - \beta_o(0))}{\alpha(0) - \beta_i(0) - \beta_o(0)}.$$

Since $z(\eta) \in C^\infty[1, \infty)$ and the arguments $\frac{\gamma(\alpha - \beta_i - \beta_o)}{(\alpha - \beta_i)(\alpha - \beta_o)}$, $\frac{\gamma}{\alpha - \beta_i}$, $\frac{\gamma}{\alpha - \beta_o}$ are bounded away from 1 for $(\alpha, \beta_i, \beta_o, \gamma) \in \mathbf{W}_\epsilon$, we have that f_1 and f_2 are smooth on \mathbf{W}_ϵ . \square

As an illustration, and for a partial check, let us evaluate the values of f_1 and f_2 for the following initial state

$$\mathbf{w}_0 = (1 - e^{-2c}, e^{-c}(1 - e^{-c}), e^{-c}(1 - e^{-c}), c).$$

We will later see that the likely initial states arising from $D(n, m = c_n n)$ are near $n\mathbf{w}_0$. We have that

$$z_{i,o}(0) = z \left(\frac{c}{1 - e^{-c}} \right) = c,$$

and

$$z^* = z \left(c \frac{(1 - e^{-2c} - e^{-c}(1 - e^{-c})^2)}{1 - e^{-2c} - 2e^{-c}(1 - e^{-c})} \right) = z(c).$$

In particular, $z^* = z(c)$ is the unique root of $\ell(z) = c$, or equivalently of the equation

$$1 - \frac{z}{c} = e^{-z} = e^{-c(z/c)},$$

so that $z(c) = c\theta(c)$. Therefore

$$\begin{aligned} f_1(\mathbf{w}_0) &= \frac{(z^*)^2}{z_i(0)z_o(0)} \cdot \frac{(1 - e^{-2c} - e^{-c}(1 - e^{-c}))(1 - e^{-2c} - e^{-c}(1 - e^{-c}))}{1 - e^{-2c} - e^{-c}(1 - e^{-c}) - e^{-c}(1 - e^{-c})} \\ &= \frac{z(c)^2}{c^2} \cdot 1 = \theta(c)^2, \end{aligned}$$

in agreement with the results by Karp [18] that the expected number of vertices in the strong giant component is approximately $\theta(c)^2 n$. Furthermore

$$f_2(\bar{\mathbf{w}}) = \gamma(0) \frac{(z^*)^2}{z_i(0)z_o(0)} = c\theta^2,$$

which means that the average in-degree and out-degree of the terminal digraph should be asymptotic to $c_n \sim c$, the original average in/out-degree of the initial digraph $D(n, m = c_n n)$.

5.4 The PDE for the covariance parameters

Turn to the covariance parameters $\{\psi_{j,k}\}$, a solution of PDE-type equation (4.8) meeting boundary conditions (4.3). The equation is

$$\hat{E}_{\mathbf{s}}[\Delta \mathbf{s}^T] \nabla \psi_{j,k} + \hat{E}_{\mathbf{s}}[(\Delta \mathbf{s}^T \nabla f_j)(\Delta \mathbf{s}^T \nabla f_k)] = o(n^{-1/2}),$$

where $\Delta \mathbf{s}^T = (-a - b, r_i - a, r_o - b, -k)$. Just as for the PDE-type equation for f_1, f_2 (4.7), it suffices to find $\psi_{j,k}$ that satisfy the corresponding PDE with $\mathcal{E}_{\mathbf{s}}[\cdot]$ instead of $\hat{E}_{\mathbf{s}}$, and the $o(n^{-1/2})$ term set equal 0. Thus, we need to solve a scale-free equation

$$\mathcal{E}_{\mathbf{s}}[\Delta \mathbf{s}^T] \nabla \psi_{j,k} + \mathcal{E}_{\mathbf{s}}[(\Delta \mathbf{s}^T \nabla f_j)(\Delta \mathbf{s}^T \nabla f_k)] = 0, \quad (5.15)$$

subject to boundary conditions: for $\alpha, \gamma > 0$, $\psi_{j,k}(\alpha, 0, 0, \gamma) = 0$.

Notice that the linear operator $\mathcal{E}_s[\Delta \mathbf{s}^T] \nabla$ is exactly the same as in the PDE for f_1 and f_2 . In particular, we have the same characteristics. However, unlike f_j , $\psi_{j,k}$ is not constant along trajectories. Instead,

$$\frac{d\psi_{j,k}}{dt}(\mathbf{w}) = -\mathcal{E}_s[(\Delta \mathbf{s}^T \nabla f_j)(\Delta \mathbf{s}^T \nabla f_k)],$$

with the non-zero RHS, expressed through the already known f_1 and f_2 .

For $\mathbf{w}(0) \in \mathbf{W}_\epsilon$, we found that $\mathbf{w}(\mathcal{T}) = (\alpha(\mathcal{T}), 0, 0, \gamma(\mathcal{T}))$, where $\alpha(\mathcal{T}), \gamma(\mathcal{T}) > 0$. So, by the boundary condition on ψ , we have that $\psi_{j,k}(\alpha(\mathcal{T}), 0, 0, \gamma(\mathcal{T})) = 0$. Therefore

$$\begin{aligned} \psi_{j,k}(\mathbf{w}(0)) &= \int_0^\mathcal{T} -\frac{d\psi_{j,k}}{dt} dt = \int_0^\mathcal{T} \Psi_{j,k}(\mathbf{w}(t)) dt, \\ \Psi_{j,k}(\mathbf{w}(t)) &:= \mathcal{E}_s[(\Delta \mathbf{s}^T \nabla f_j)(\Delta \mathbf{s}^T \nabla f_k)]. \end{aligned} \quad (5.16)$$

Observe that, for $\mathbf{u} \in \mathbb{R}^2$, and all $t \in [0, \mathcal{T}]$,

$$\mathbf{u}^T \begin{pmatrix} \Psi_{1,1} & \Psi_{1,2} \\ \Psi_{2,1} & \Psi_{2,2} \end{pmatrix} \mathbf{u} = \mathcal{E}_s \left[\left(\sum_{j=1}^2 u_j \Delta \mathbf{s}^T \nabla f_j \right)^2 \right] \geq 0.$$

Consequently, without actually evaluating the integral in (5.16), we already know that

$$\mathbf{u}^T \begin{pmatrix} \psi_{1,1}(\mathbf{w}(0)) & \psi_{1,2}(\mathbf{w}(0)) \\ \psi_{2,1}(\mathbf{w}(0)) & \psi_{2,2}(\mathbf{w}(0)) \end{pmatrix} \mathbf{u} \geq 0.$$

Of course, this inequality should hold since $n\{\psi_{j,k}(\mathbf{w}(0))\}_{1 \leq j,k \leq 2}$ is anticipated to be the limiting covariance matrix for the initial state $\mathbf{s}(0) = n\mathbf{w}(0)$. A drawback of (5.16) is that \mathcal{T} is defined implicitly by $\beta_i(\mathcal{T}) = \beta_o(\mathcal{T}) = 0$. Using a change of variables, namely switching from t say to z_i and using our formula for $\frac{dz_i}{dt}$ (see (5.7)), we replace (5.16) by

$$\begin{aligned} \psi_{j,k}(\mathbf{w}(0)) &= \int_{z^*(0)}^{z_i(0)} \Psi_{j,k}(\mathbf{w}) \frac{(\beta_i + \beta_o)(\alpha - \beta_o)}{\beta_i z_i} dz_i, \\ z^*(0) &= z \left(\frac{\gamma(0)(\alpha(0) - \beta_i(0) - \beta_o(0))}{(\alpha(0) - \beta_i(0))(\alpha(0) - \beta_o(0))} \right). \end{aligned} \quad (5.17)$$

Indeed, here $\mathbf{w} = \mathbf{w}(z_i) = (\alpha, \beta_i, \beta_o, \gamma)$ is the solution of the ODE (5.6) where we switch from t to z_i , i.e.

$$\frac{d\mathbf{w}}{dz_i} = \frac{(\beta_i + \beta_o)(\alpha - \beta_o)}{-\beta_i z_i} \mathbf{H}(\mathbf{w}); \quad \mathbf{w}(z_i(0)) = \mathbf{w}(0).$$

Let's sketch the derivation of formulas for $\Psi_{j,k}(\mathbf{w})$. Since both $\Delta \mathbf{s}$ and ∇f_j are 4-dimensional, there are 16 summands in $\mathcal{E}_s[(\Delta \mathbf{s}^T \nabla f_j)(\Delta \mathbf{s}^T \nabla f_k)]$ for each (j, k) . Indeed, denoting $f_x = \partial f / \partial x$,

$$\begin{aligned}
(\Delta \mathbf{s}^T \nabla f_j)(\Delta \mathbf{s}^T \nabla f_k) &= (a+b)^2(f_j)_\alpha(f_k)_\alpha + (r_i-a)(-a-b)(f_j)_{\beta_i}(f_k)_\alpha \\
&+ (r_o-b)(-a-b)(f_j)_{\beta_o}(f_k)_\alpha + k(a+b)(f_j)_\gamma(f_k)_\alpha + (-a-b)(r_i-a)(f_j)_\alpha(f_k)_{\beta_i} \\
&+ (r_i-a)^2(f_j)_{\beta_i}(f_k)_{\beta_i} + (r_o-b)(r_i-a)(f_j)_{\beta_o}(f_k)_{\beta_i} - k(r_i-a)(f_j)_\gamma(f_k)_{\beta_i} \quad (5.18) \\
&+ (-a-b)(r_o-b)(f_j)_\alpha(f_k)_{\beta_o} + (r_i-a)(r_o-b)(f_j)_{\beta_i}(f_k)_{\beta_o} + (r_o-b)^2(f_j)_{\beta_o}(f_k)_{\beta_o} \\
&- k(r_o-b)(f_j)_\gamma(f_k)_{\beta_o} + (-a-b)(-k)(f_j)_\alpha(f_k)_\gamma + (r_i-a)(-k)(f_j)_{\beta_i}(f_k)_\gamma \\
&+ (r_o-b)(-k)(f_j)_{\beta_o}(f_k)_\gamma - k(-k)(f_j)_\gamma(f_k)_\gamma.
\end{aligned}$$

Clearly, this is a sum of quadratic polynomials of the random variables a, b, r_i, r_o, k (i.e. $(a+b)^2, \dots, k^2$) with deterministic coefficients being pairwise products of partial derivatives of f_1 and f_2 taken at the current $\mathbf{w} = \mathbf{w}(t)$. The \mathcal{E}_s expected values of these polynomials can be easily obtained from the 15 \mathcal{E}_s expected values (of a^2, ab, \dots, k^2) given in Appendix A.1. As for the gradient of f_j (given in Proposition 5.7), they are computed by using implicit differentiation on $z(\circ)$. We refer the dedicated reader to a Mathematica notebook file, which contains all of these terms at <http://www.dpoole.info/strong-giant/>.

Since f_1, f_2 are smooth on \mathbf{W}_ϵ , then so are the resulting $\Psi_{j,k}$. Therefore, the integrals in (5.17) are well-defined, which proves existence of the sought-after solution of the PDE for $\{\psi_{j,k}\}$. The expressions for the integrands $\Psi_{j,k}(\mathbf{w})$ are exceedingly long and attempts to simplify them were unsuccessful. However, it follows from smoothness of f_1, f_2 and all $\Psi_{j,k}$ that the functions $\psi_{j,k}(\mathbf{w})$ are smooth as well!

In summary, we have found smooth $f_j(\mathbf{w})$ and proved existence of smooth $\psi_{j,k}(\mathbf{w})$, the solutions of the partial differential equations (5.2) and (5.15).

Getting back to (4.7)–(4.8), we have proved that, uniformly over $\mathbf{s} \in \mathbf{S}_\epsilon$,

$$\hat{E}_s[\Delta \mathbf{s}^T] \nabla f_j(\mathbf{s}/n) = O((\ln n)^8/n),$$

and

$$\hat{E}_s[\Delta \mathbf{s}^T] \nabla \psi_{j,k} + \hat{E}_s[(\Delta \mathbf{s}^T \nabla f_j)(\Delta \mathbf{s}^T \nabla f_k)] = O((\ln n)^8/n).$$

Consequently, uniformly over $\|\mathbf{u}\| = O(n^{-1/2})$ and $\mathbf{s} \in \mathbf{S}_\epsilon$,

$$G_n(\mathbf{s}/n, \mathbf{u}) - \sum_{\mathbf{s}'} G_n(\mathbf{s}'/n, \mathbf{u}) \hat{P}(\mathbf{s}(t+1) = \mathbf{s}' | \mathbf{s}(t) = \mathbf{s}) = O((\ln n)^8/n^{3/2}), \quad (5.19)$$

cf. (4.6).

Thus we have found the Gaussian characteristic function G_n very nearly satisfying (5.19) for $\mathbf{s} \in \mathbf{S}_\epsilon$, the equation satisfied perfectly by the actual characteristic function $\hat{\varphi}$ for all \mathbf{s} . Our next step is to show that, for $\mathbf{s}(0) \in \mathbf{S}_{\epsilon'}$, with ϵ' close to ϵ from below, a.a.s. the process $\{\mathbf{s}(t)\}$ leaves \mathbf{S}_ϵ simply because the number of semi-isolated vertices, $\nu_i + \nu_o$, drops down to zero; formally, a.a.s. $\hat{\tau} = \bar{\tau}$. That is, establishing the asymptotic normality of $(\hat{\nu}, \hat{\mu})$ is all we need for asymptotic normality of $(\bar{\nu}, \bar{\mu})$.

6 Likely reason for leaving \mathbf{S}_ϵ

So we want to show that for $\mathbf{s}(0) \in \mathbf{S}_{\epsilon'}$, with properly chosen $\epsilon' < \epsilon$, a.a.s. $\mathbf{s}(t)$ leaves \mathbf{S}_ϵ at the first moment t when $\nu_i(t) + \nu_o(t) = 0$. To do so, we will use constancy of $I_j(\mathbf{w}(t))$ along the characteristic $\{\mathbf{w}(t)\}$ to show that their counterparts $F_1(\mathbf{s}) = \mu(\nu - \nu_i - \nu_o)/[(\nu - \nu_i)(\nu - \nu_o)]$ and $F_2(\mathbf{s}) = z_i z_o / [\mu/n]$ a.a.s. are almost constant along the random process $\{\mathbf{s}(t)\}$, for as long as $\nu_i(t) + \nu_o(t) > 0$. For the proof, we use the integrals $I_j(\mathbf{w})$ to construct a pair of exponential supermartingales, and then apply the maximum inequality for supermartingales based on the Optional Sampling Theorem (Durrett [13]). This approach had been used in [33] and in [1] for the deletion processes for k -core and Karp-Sisper greedy matching problems, respectively.

For a given $L = L_n$, define

$$Q_j(\mathbf{s}) := \exp \left(L(F_j(\mathbf{s}) - F_j(\mathbf{s}(0))) \right).$$

Consider the process $Q^j(t) := Q_j(\mathbf{s}(t))1_{\{t < \hat{\tau}\}}$, for $t \geq 0$.

Lemma 6.1. *If $L_n = o((\ln n)(\ln \ln n))$, then the process*

$$R^j(t) := (1 + n^{-1})^{-t} Q^j(t) \tag{6.1}$$

is a non-negative supermartingale for $j = 1, 2$.

Proof. There is little difference between the proofs for $j = 1$ and $j = 2$, so we consider $j = 1$ only. It suffices to show that

$$\hat{E}[Q^1(t+1) | \{\mathbf{s}(t')\}_{t' \leq t}] \leq Q^1(t)(1 + n^{-1}), \tag{6.2}$$

for all t , since (6.2) can be rewritten as

$$\hat{E}[R^1(t+1) | \{\mathbf{s}(t')\}_{t' \leq t}] \leq R^1(t).$$

If $\mathbf{s}(t) \notin \mathbf{S}_\epsilon$, then $Q^1(t+1) = Q^1(t)$ (see definition of $\hat{\tau}$ in (4.4)), so that

$$\hat{E}[Q^1(t+1) | \{\mathbf{s}(t')\}_{t' \leq t}] = Q^1(t), \tag{6.3}$$

and obviously (6.2) holds. Suppose $\mathbf{s}(t) \in \mathbf{S}_\epsilon$. In this case, the (conditional) expectation with respect to \hat{P} is the expectation with respect to P . Using the definition of $Q^1(\cdot)$, we obtain

$$\begin{aligned} E[Q^1(\mathbf{s}(t+1)) | \{\mathbf{s}(t')\}_{t' \leq t}] &= Q^1(t) \sum_{\mathbf{s}' \in \mathbf{S}_\epsilon} e^{L(F_1(\mathbf{s}') - F_1(\mathbf{s}(t)))} P(\mathbf{s}(t+1) = \mathbf{s}' | \mathbf{s}(t)) \\ &\leq Q^1(t)(E_1 + E_2); \\ E_1 &:= \sum_{\mathbf{s}': k \leq \ln n} e^{L(F_1(\mathbf{s}') - F_1(\mathbf{s}(t)))} P(\mathbf{s}(t+1) = \mathbf{s}' | \mathbf{s}(t)), \\ E_2 &:= \sum_{\substack{\mathbf{s}' \in \mathbf{S}_\epsilon \\ k > \ln n}} e^{L(F_1(\mathbf{s}') - F_1(\mathbf{s}(t)))} P(\mathbf{s}(t+1) = \mathbf{s}' | \mathbf{s}(t)). \end{aligned}$$

Consider E_1 . Using Fact 3.7, uniformly over $\mathbf{s}(t) \in \mathbf{S}_\epsilon$ and \mathbf{s}' such that $k \leq \ln n$,

$$\begin{aligned} F_1(\mathbf{s}') &= \frac{\mu'(\nu' - \nu'_i - \nu'_o)}{(\nu' - \nu'_i)(\nu' - \nu'_o)} = \frac{\mu(\nu - \nu_i - \nu_o)}{(\nu - \nu_i)(\nu - \nu_o)} + O(n^{-1} \ln n) \\ &= F_1(\mathbf{s}(t)) + O(n^{-1} \ln n), \end{aligned}$$

and so

$$L(F_1(\mathbf{s}') - F_1(\mathbf{s}(t))) = O\left(\frac{(\ln n)^2 \ln \ln n}{n}\right) \rightarrow 0.$$

Therefore

$$\exp(L(F_1(\mathbf{s}') - F_1(\mathbf{s}(t)))) = 1 + L(F_1(\mathbf{s}') - F_1(\mathbf{s}(t))) + O((\ln n)^5/n^2).$$

Again using Fact 3.7, it is easy to check that $\|\nabla(F_1(\mathbf{s}(t)))\| = O(n^{-1})$ and that all 6 second-order partial derivatives of $F_1(\mathbf{s})$ at $\mathbf{s} = (1 - \lambda)\mathbf{s}(t) + \lambda\mathbf{s}'$, $0 \leq \lambda \leq 1$, are of order $O(n^{-2})$, uniformly for $\mathbf{s}(t) \in \mathbf{S}_\epsilon$, \mathbf{s}' and λ , if $k \leq \ln n$. So, using $\|\Delta \mathbf{s}\| = \|\mathbf{s}' - \mathbf{s}(t)\| = O(\ln n)$, we have

$$F_1(\mathbf{s}') - F_1(\mathbf{s}(t)) = \Delta \mathbf{s}^T (\nabla F_1(\mathbf{s}(t))) + O((\ln n)^2/n^2) = O((\ln n)/n),$$

Consequently

$$\begin{aligned} E_1 &= \sum_{\mathbf{s}': k \leq \ln n} (1 + L\Delta \mathbf{s}^T \nabla F_1(\mathbf{s}(t)) + O((\ln n)^5/n^2)) P(\Delta \mathbf{s} | \mathbf{s}(t)) \\ &= P(k \leq \ln n | \mathbf{s}(t)) + O((\ln n)^5/n^2) + L \sum_{\mathbf{s}': k \leq \ln n} \Delta \mathbf{s}^T (\nabla F_1(\mathbf{s}(t))) P(\Delta \mathbf{s} | \mathbf{s}(t)). \end{aligned}$$

Now we found earlier, (5.1), that uniformly over $\mathbf{s}(t) \in \mathbf{S}_\epsilon$,

$$\sum_{\mathbf{s}': k \leq \ln n} \Delta \mathbf{s}^T P(\Delta \mathbf{s} | \mathbf{s}(t)) = \mathcal{E}_{\mathbf{s}(t)}[\Delta \mathbf{s}^T] + O((\ln n)^8/n).$$

Also, by (5.9), since $I_j(\mathbf{w}(t))$ is constant along the trajectory $d\mathbf{w}/dt = \mathcal{E}_{\mathbf{s}(t)}[\Delta \mathbf{s}]$, the gradient $\nabla I_1(\mathbf{s}/n)$ is orthogonal to $\mathcal{E}_{\mathbf{s}(t)}[\Delta \mathbf{s}]$, or equivalently $\nabla F_1(\mathbf{s}) \cdot \mathcal{E}_{\mathbf{s}(t)}[\Delta \mathbf{s}] = 0$. Combining this with the estimate $L\|\nabla F_1(\mathbf{s}(t))\| = O((\ln n)^2/n)$ (uniformly over $\mathbf{s}(t) \in \mathbf{S}_\epsilon$), we have that

$$L \sum_{\mathbf{s}': k \leq \ln n} \Delta \mathbf{s}^T \nabla F_1(\mathbf{s}(t)) P(\Delta \mathbf{s} | \mathbf{s}(t)) = O\left(L\|\nabla F_1\| \frac{(\ln n)^8}{n}\right) \ll \frac{(\ln n)^{10}}{n^2}.$$

Therefore

$$E_1 = P(k \leq \ln n | \mathbf{s}(t)) + O((\ln n)^{10}/n^2).$$

Here, by Lemma 3.9, (ii), we have that

$$P(k > \ln n | \mathbf{s}(t)) = \sum_{\mathbf{s}': k > \ln n} P(\mathbf{s}(t+1) = \mathbf{s}' | \mathbf{s}(t)) \leq e^{-\frac{2}{3}(\ln n)(\ln \ln n)}, \quad (6.4)$$

which implies that

$$E_1 = 1 + O((\ln n)^{10}/n^2). \quad (6.5)$$

Consider E_2 . Since $F_1(\mathbf{s}') \leq c_n + \epsilon$ for $\mathbf{s}' \in \mathbf{S}_\epsilon$, we have that

$$L(F_1(\mathbf{s}') - F_1(\mathbf{s}(0))) \leq L(c_n + \epsilon) = o((\ln n)(\ln \ln n)).$$

Using the bound (6.4), we have that

$$\sum_{\mathbf{s}': k > \ln n} e^{L(F_1(\mathbf{s}') - F_1(\mathbf{s}(0)))} P(\mathbf{s}(t+1) = \mathbf{s}' | \mathbf{s}(t)) = O\left(e^{-\frac{2}{3}(\ln n)(\ln \ln n)}\right) \ll \frac{(\ln n)^{10}}{n^2}. \quad (6.6)$$

Combining the bounds on E_1 , (6.5), and E_2 , (6.6), we obtain that for $\mathbf{s}(t) \in \mathbf{S}_\epsilon$,

$$E[1_{\{t+1 < \hat{\tau}\}} e^{L(F_1(\mathbf{s}(t+1)) - F_1(\mathbf{s}(t)))}] \leq 1 + O((\ln n)^{10}/n^2) < 1 + n^{-1}. \quad (6.7)$$

Hence, by (6.3) and (6.7), for any $\mathbf{s}(t)$,

$$\hat{E}[Q^1(t+1) | \mathbf{s}(t)] \leq Q^1(t) (1 + n^{-1}),$$

and $\{R^1(t)\}$, defined in (6.1), is a non-negative supermartingale. \square

Now we are ready to prove that $F_j(\mathbf{s}(t))$ is almost constant all the way to (and especially including) the time $\hat{\tau}$. To state the result and for future usage, we borrow a term from Knuth, Motwani and Pittel [20]: we say that an event, A , holds *quite surely* (*q.s.* in short) if for each fixed $a > 0$,

$$P(A) \geq 1 - n^{-a},$$

for n sufficiently large.

Lemma 6.2. *Let $L_n = (\ln n)(\ln \ln n) / \ln \ln \ln n$. Uniformly over $\mathbf{s}(0) \in \mathbf{S}_\epsilon$, q.s.*

$$\max_{t \leq \hat{\tau}} |F_j(\mathbf{s}(t)) - F_j(\mathbf{s}(0))| \leq \rho_n := \frac{1}{\ln \ln \ln n}. \quad (6.8)$$

Proof. Introduce yet another stopping time

$$\tau' = \begin{cases} \min\{t \leq \hat{\tau} : F_j(\mathbf{s}(t)) - F_j(\mathbf{s}(0)) > \rho_n\}, \\ \hat{\tau} + 1, & \text{if no such } t \text{ exists.} \end{cases}$$

Applying the Optional Sampling Theorem to the supermartingale $R^j(t)$ and the stopping time τ' (see Durrett [13]), we have that

$$\begin{aligned} E[Q^j(\tau')] &= E[(1 + n^{-1})^{\tau'} R^j(\tau')] \leq (1 + n^{-1})^n E[R^j(0)] \\ &= (1 + n^{-1})^n E[Q^j(0)] = (1 + n^{-1})^n \leq e. \end{aligned}$$

On the event $\{\tau' \leq \hat{\tau}\}$, we have that $Q^j(\tau') \geq e^{L\rho_n}$ and so

$$e^{L\rho_n} P(\tau' \leq \hat{\tau}) \leq E[Q^j(\tau')] \leq e.$$

Equivalently,

$$\begin{aligned} P\left(\max_{t \leq \hat{\tau}} [F_j(\mathbf{s}(t)) - F_j(\mathbf{s}(0))] > \rho_n\right) &= P(\tau' \leq \hat{\tau}) = O(e^{-L\rho_n}) \\ &= O\left(e^{-(\ln n) \frac{\ln \ln n}{(\ln \ln \ln n)^2}}\right) \ll n^{-a}, \quad \forall \text{ (fixed) } a > 0. \end{aligned}$$

The case $F_j(\mathbf{s}(t)) - F_j(\mathbf{s}(0)) \leq -\rho_n$ is treated similarly. \square

Finally,

Lemma 6.3. *Suppose $\epsilon' > 0$ is such that $\epsilon' + \rho_n \leq \epsilon$, where ρ_n is defined in (6.8). Uniformly over $\mathbf{s}(0) \in \mathbf{S}_{\epsilon'}$, q.s. $F_j(\mathbf{s}(\hat{\tau})) \in (c_n - \epsilon, c_n + \epsilon)$, whence q.s. $\nu_i(\hat{\tau}) + \nu_o(\hat{\tau}) = 0$. Thus q.s. $\mathbf{s}(t)$ leaves \mathbf{S}_{ϵ} because the number of semi-isolated vertices drops down to 0, not because $F_j(\mathbf{s}(t))$ escapes $(c_n - \epsilon, c_n + \epsilon)$. In other words, q.s. $\hat{\tau} = \bar{\tau}$.*

Proof. By the definition of $\mathbf{S}_{\epsilon'}$ (3.9), we have $F_j(\mathbf{s}(0)) \in (c_n - \epsilon', c_n + \epsilon')$, and by Lemma 6.2, uniformly over $\mathbf{s}(0) \in \mathbf{S}_{\epsilon'}$, q.s. $|F_j(\mathbf{s}(\hat{\tau})) - F_j(\mathbf{s}(0))| \leq \rho_n$. Therefore q. s.

$$F_j(\mathbf{s}(\hat{\tau})) \in (c_n - \epsilon' - \rho_n, c_n + \epsilon' + \rho_n) \subset (c_n - \epsilon, c_n + \epsilon).$$

However, if $F_{1,2}(\mathbf{s}(\hat{\tau})) \in (c_n - \epsilon, c_n + \epsilon)$, then— by the definition of \mathbf{S}_{ϵ} — the only remaining reason for the process to leave \mathbf{S}_{ϵ} is that at this moment $\hat{\tau}$ the total number of semi-isolated vertices has fallen to zero. \square

Clearly, if $(\hat{\nu}, \hat{\mu})$ is shown to be asymptotically Gaussian, then so will be $(\bar{\nu}, \bar{\mu})$.

7 Distribution of $(\hat{\nu}, \hat{\mu})$ starting from a generic \mathbf{s}

In this section, we prove that indeed, with $\mathbf{s}(0) \in \mathbf{S}_{\epsilon'}$, for a certain ϵ' , the terminal pair $(\hat{\nu}, \hat{\mu})$, once centered and scaled, is asymptotically Gaussian. To do so we will show that the characteristic function, $\hat{\varphi}_{\mathbf{s}}(\mathbf{u})$, of the pair $(\hat{\nu}, \hat{\mu})$ is closely approximated by the Gaussian characteristic function $G_n(\mathbf{s}(0)/n, \mathbf{u})$ (defined in (4.2)). We begin with noticing that $\hat{\varphi}_{\mathbf{s}}(\mathbf{u})$ satisfies the one-step recurrence relation

$$\hat{\varphi}_{\mathbf{s}}(\mathbf{u}) - \sum_{\mathbf{s}^{(1)}} \hat{\varphi}_{\mathbf{s}^{(1)}}(\mathbf{u}) \hat{P}(\mathbf{s}^{(1)} | \mathbf{s}) = 0.$$

Iterating this recurrence n times, n being the largest possible number of deletion steps, we find that

$$\hat{\varphi}_{\mathbf{s}}(\mathbf{u}) - \sum_{\mathbf{s}^{(1)} \dots \mathbf{s}^{(n)}} \hat{\varphi}_{\mathbf{s}^{(n)}}(\mathbf{u}) \prod_{j=1}^n \hat{P}(\mathbf{s}^{(j)} | \mathbf{s}^{(j-1)}) = 0, \quad \mathbf{s}^{(0)} = \mathbf{s}. \quad (7.1)$$

Let us show that $G_n(\mathbf{s}/n, \mathbf{u})$ satisfies (7.1) with a $o(1)$ error term replacing 0 on the right.

Proposition 7.1. *Uniformly for $\|\mathbf{u}\| = O(n^{-1/2})$,*

$$\left| G_n(\mathbf{s}/n, \mathbf{u}) - \sum_{\mathbf{s}^{(1)} \dots \mathbf{s}^{(n)}} G_n(\mathbf{s}^{(n)}/n, \mathbf{u}) \hat{P}(\{\mathbf{s}(t)\} = \{\mathbf{s}^{(i)}\}) \right| \leq n^{-0.49}. \quad (7.2)$$

Proof. Define

$$D_n(\mathbf{s}/n, \mathbf{u}) := G_n(\mathbf{s}/n, \mathbf{u}) - \sum_{\mathbf{s}^{(1)}} G_n(\mathbf{s}^{(1)}/n, \mathbf{u}) P(\mathbf{s}(t+1) = \mathbf{s}^{(1)} | \mathbf{s}(t) = \mathbf{s}).$$

By (5.19), uniformly over $\mathbf{s} \in \mathbf{S}_\epsilon$ and $\|\mathbf{u}\| = O(n^{-1/2})$,

$$|D_n(\mathbf{s}/n, \mathbf{u})| = O((\ln n)^8 / n^{3/2}).$$

Since \hat{P} is the transition probability for the process that freezes at the moment $\mathbf{s}(t)$ leaves \mathbf{S}_ϵ , we have: for $\mathbf{s} \notin \mathbf{S}_\epsilon$, and all t ,

$$G_n(\mathbf{s}/n, \mathbf{u}) = \sum_{\mathbf{s}^{(1)}} G_n(\mathbf{s}^{(1)}/n, \mathbf{u}) 1_{\{\mathbf{s}\}}(\mathbf{s}^{(1)}) = \sum_{\mathbf{s}^{(1)}} G_n(\mathbf{s}^{(1)}/n, \mathbf{u}) \hat{P}(\mathbf{s}(t+1) = \mathbf{s}^{(1)} | \mathbf{s}(t) = \mathbf{s}).$$

Thus, $D_n(\mathbf{s}/n, \mathbf{u}) = 0$ for $\mathbf{s} \notin \mathbf{S}_\epsilon$, whence uniformly for all \mathbf{s} and $\|\mathbf{u}\| = O(n^{-1/2})$,

$$|D_n(\mathbf{s}/n, \mathbf{u})| = O((\ln n)^8 / n^{3/2}) \leq n^{-1.49}, \quad (7.3)$$

for n large enough. It suffices to prove that for $1 \leq k \leq n$, we have that

$$\left| G_n(\mathbf{s}/n, \mathbf{u}) - \sum_{\mathbf{s}^{(1)} \dots \mathbf{s}^{(k)}} G_n(\mathbf{s}^{(k)}/n, \mathbf{u}) \prod_{j=1}^k \hat{P}(\mathbf{s}^{(j)} | \mathbf{s}^{(j-1)}) \right| \leq k \times n^{-1.49}. \quad (7.4)$$

Let us prove the above inequality by induction. The base case is precisely (7.3). Suppose that (7.4) holds for $(k-1)$. Now, we have

$$\begin{aligned} & \left| G_n(\mathbf{s}/n, \mathbf{u}) - \sum_{\mathbf{s}^{(1)} \dots \mathbf{s}^{(k)}} G_n(\mathbf{s}^{(k)}/n, \mathbf{u}) \prod_{j=1}^k \hat{P}(\mathbf{s}^{(j)} | \mathbf{s}^{(j-1)}) \right| \\ &= \left| G_n(\mathbf{s}/n, \mathbf{u}) - \sum_{\mathbf{s}^{(1)} \dots \mathbf{s}^{(k-1)}} \prod_{j=1}^{k-1} \hat{P}(\mathbf{s}^{(j)} | \mathbf{s}^{(j-1)}) \sum_{\mathbf{s}^{(k)}} G_n(\mathbf{s}^{(k)}/n, \mathbf{u}) \hat{P}(\mathbf{s}^{(k)} | \mathbf{s}^{(k-1)}) \right| \\ &= \left| G_n(\mathbf{s}/n, \mathbf{u}) - \sum_{\mathbf{s}^{(1)} \dots \mathbf{s}^{(k-1)}} \prod_{j=1}^{k-1} \hat{P}(\mathbf{s}^{(j)} | \mathbf{s}^{(j-1)}) [G_n(\mathbf{s}^{(k-1)}/n, \mathbf{u}) + D_n(\mathbf{s}^{(k-1)}/n, \mathbf{u})] \right| \\ &\leq \left| G_n(\mathbf{s}/n, \mathbf{u}) - \sum_{\mathbf{s}^{(1)} \dots \mathbf{s}^{(k-1)}} \prod_{j=1}^{k-1} \hat{P}(\mathbf{s}^{(j)} | \mathbf{s}^{(j-1)}) G_n(\mathbf{s}^{(k-1)}/n, \mathbf{u}) \right| + \max_{\mathbf{s}'} |D_n(\mathbf{s}'/n, \mathbf{u})|. \end{aligned}$$

The first term is at most $(k-1) \times n^{-1.49}$ by hypothesis and the second term is at most $n^{-1.49}$ by (7.3), which completes the proof of (7.4). \square

Next, we use this Lemma to bound the difference between $\hat{\varphi}_{\mathbf{s}}(\mathbf{u})$ and $G_n(\mathbf{s}/n, \mathbf{u})$.

Theorem 7.2. *Let $0 < \epsilon' < \epsilon$ with $\epsilon' = \epsilon'(n)$ possibly tending to zero and ϵ being fixed. For $\|\mathbf{u}\| = O(n^{-1/2})$, uniformly over $\mathbf{s}(0) \in \mathbf{S}_{\epsilon'}$,*

$$|G_n(\mathbf{s}(0)/n, \mathbf{u}) - \hat{\varphi}_{\mathbf{s}(0)}(\mathbf{u})| \leq 2n^{-.49}. \quad (7.5)$$

Proof. First, by (7.2) and (7.1), uniformly over $\|\mathbf{u}\| = O(n^{-1/2})$ and all \mathbf{s} ,

$$|G_n(\mathbf{s}/n, \mathbf{u}) - \hat{\varphi}_{\mathbf{s}}(\mathbf{u})| \leq \sum_{\mathbf{s}^{(1)}, \dots, \mathbf{s}^{(n)}} |G_n(\mathbf{s}^{(n)}/n, \mathbf{u}) - \hat{\varphi}_{\mathbf{s}^{(n)}}(\mathbf{u})| \hat{P}(\{\mathbf{s}(t)\} = \{\mathbf{s}^{(i)}\}) + n^{-.49}. \quad (7.6)$$

Here $\mathbf{s}^{(n)} = \mathbf{s}(\hat{\tau})$, of course. So, by Lemma 6.3, uniformly over $\mathbf{s}(0) \in \mathbf{S}_{\epsilon'}$, with probability at least $1 - n^{-1}$, we have $\mathbf{s}^{(n)} \in \mathbf{S}$, where

$$\mathbf{S} = \{\mathbf{s} = (\nu, 0, 0, \mu) : \nu, \mu > 0, F_1(\mathbf{s}), F_2(\mathbf{s}) \in (c_n - \epsilon, c_n + \epsilon)\}.$$

Accordingly, we break up the sum in (7.6) into two parts, over $\mathbf{s}^{(n)} \in \mathbf{S}$ and over $\mathbf{s} \notin \mathbf{S}$. Using the obvious $|G_n(\mathbf{s}^{(n)}/n, \mathbf{u}) - \hat{\varphi}_{\mathbf{s}^{(n)}}(\mathbf{u})| \leq 2$ for the second sum, we obtain

$$\begin{aligned} |G_n(\mathbf{s}/n, \mathbf{u}) - \hat{\varphi}_{\mathbf{s}}(\mathbf{u})| &\leq \sum_{\substack{\mathbf{s}^{(1)}, \dots, \mathbf{s}^{(n-1)}, \\ \mathbf{s}^{(n)} \in \mathbf{S}}} |G_n(\mathbf{s}^{(n)}/n, \mathbf{u}) - \hat{\varphi}_{\mathbf{s}^{(n)}}(\mathbf{u})| \hat{P}(\{\mathbf{s}(t)\} = \{\mathbf{s}^{(i)}\}) \\ &\quad + 2 \sum_{\substack{\mathbf{s}^{(1)}, \dots, \mathbf{s}^{(n-1)}, \\ \mathbf{s}^{(n)} \notin \mathbf{S}}} \hat{P}(\{\mathbf{s}(t)\} = \{\mathbf{s}^{(i)}\}) + n^{-0.49}. \end{aligned} \quad (7.7)$$

This first sum is zero. Here is why: by the definition of \mathbf{S} , if $\mathbf{s} \in \mathbf{S}$, then $\mathbf{s} \notin \mathbf{S}_{\epsilon}$ (defined in (3.9)), because $\nu_i = \nu_o = 0$. So the process that starts at $\mathbf{s} \notin \mathbf{S}_{\epsilon}$ stays at \mathbf{s} . So, for any $\mathbf{s} \in \mathbf{S}$,

$$\hat{\varphi}_{\mathbf{s}}(\mathbf{u}) = \hat{E}_{\mathbf{s}}[\exp(iu_1\hat{\nu} + iu_2\hat{\mu})] = \exp(iu_1\nu + iu_2\mu),$$

and, by the boundary conditions on $f_j(\mathbf{w})$, $\psi_{j,k}(\mathbf{w})$,

$$f_1(\mathbf{s}/n) = \nu/n, \quad f_2(\mathbf{s}/n) = \mu/n; \quad \psi_{j,k}(\mathbf{s}/n) = 0,$$

whence

$$G(\mathbf{s}/n, \mathbf{u}) = \exp(iu_1\nu + iu_2\mu),$$

as well. The second sum in (7.7) is precisely $P(\mathbf{s}(\hat{\tau}) \notin \mathbf{S})$, just proved to be at most n^{-1} . So (7.5) follows. \square

Corollary 7.3. *Let $\epsilon' \rightarrow 0$ however slowly. For $\mathbf{s} \in \mathbf{S}_{\epsilon'}$,*

$$\left(\frac{\bar{\nu} - nf_1(\mathbf{s}/n)}{\sqrt{n}}, \frac{\bar{\mu} - nf_2(\mathbf{s}/n)}{\sqrt{n}} \right) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, \boldsymbol{\psi}), \quad (7.8)$$

where $\mathcal{N}(\mathbf{0}, \boldsymbol{\psi})$ is a Gaussian distribution with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\psi} = (\psi_{j,k}(\mathbf{s}/n))$.

Proof. Fix $\mathbf{t} \in \mathbb{R}^2$ and set $\mathbf{u} = n^{-1/2}\mathbf{t}$. By Theorem 7.2,

$$E_{\mathbf{s}} \left[e^{i\mathbf{t}^T \frac{(\hat{\nu}, \hat{\mu})^T}{\sqrt{n}}} \right] - \exp \left(i\sqrt{n}\mathbf{t}^T \mathbf{f}(\mathbf{s}/n) - \frac{1}{2}\mathbf{t}^T \boldsymbol{\psi}(\mathbf{s}/n)\mathbf{t} \right) = O(n^{-0.49}).$$

Equivalently,

$$E_{\mathbf{s}} \left[e^{i\mathbf{t}^T \frac{(\hat{\nu}, \hat{\mu})^T - n\mathbf{f}(\mathbf{s}/n)}{\sqrt{n}}} \right] - \exp \left(-\frac{1}{2}\mathbf{t}^T \boldsymbol{\psi}(\mathbf{s}/n)\mathbf{t} \right) = O(n^{-0.49}).$$

Now, since q.s. $\hat{\tau} = \bar{\tau}$, we may replace $(\hat{\nu}, \hat{\mu})$ in the above expectation with $(\bar{\nu}, \bar{\mu})$ along with an additive error of order at most n^{-a} for any fixed $a > 0$. Choosing $a > 0.49$, we have that

$$E_{\mathbf{s}} \left[e^{i\mathbf{t}^T \frac{(\bar{\nu}, \bar{\mu})^T - n\mathbf{f}(\mathbf{s}/n)}{\sqrt{n}}} \right] - \exp \left(-\frac{1}{2}\mathbf{t}^T \boldsymbol{\psi}(\mathbf{s}/n)\mathbf{t} \right) = O(n^{-0.49}). \quad (7.9)$$

By the Multivariate Continuity Theorem, (see Durrett [13]), we have then (7.8). \square

We are close now to proving a similar result for the deletion process applied to a random instance of $\mathbf{s}(0)$ induced by $D(n, m = c_n n)$. To achieve this, in the next section we prove that $\mathbf{s}(0)$ is asymptotically Gaussian with a certain mean and a certain covariance matrix, both linear in n . This implies that the components of $\mathbf{s}(0)/n$, whence $f_j(\mathbf{s}(0)/n)$, $\psi(\mathbf{s}(0)/n)$, experience random fluctuations of magnitude $n^{-1/2}$. Combining this information with Corollary 7.3, we will see that $(\bar{\nu}, \bar{\mu})$ for the random $D(n, m = c_n n)$ is asymptotically Gaussian with mean computed for $\mathbf{s}^* := E[\mathbf{s}(0)/n]$ and the 2×2 covariance matrix obtained from $n\psi_{j,k}(\mathbf{s}^*/n)$ by adding a certain non-negative 2×2 matrix.

8 Asymptotic Distribution of $\mathbf{s}(0)$ in $D(n, m = c_n n)$

First of all, we write $\mathbf{s}(0) = (X, X_i, X_o, c_n n)$; here X, X_i, X_o are the number of non-isolated vertices, the number of vertices with zero in-degree and non-zero out-degree, and the number of vertices with zero out-degree and non-zero in-degree in $D(n, m = c_n n)$.

Lemma 8.1. *Let $c_n \rightarrow c \in (1, \infty)$. Then*

$$\left(\frac{X - (1 - e^{-2c_n})n}{\sqrt{n}}, \frac{X_i - (e^{-c_n} - e^{-2c_n})n}{\sqrt{n}}, \frac{X_o - (e^{-c_n} - e^{-2c_n})n}{\sqrt{n}} \right) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, \mathbf{K})$$

where $\mathcal{N}(\mathbf{0}, \mathbf{K})$ is a 3-dimensional Gaussian distribution with $\mathbf{0}$ and a covariance matrix $\mathbf{K} = \mathbf{K}(c)$.

Comment 8.2. *The following proof actually works for $c \in (0, \infty)$, but to avoid conflicting with the convention thus far that $c_n \rightarrow c \in (1, \infty)$, we restrict the range.*

Proof. For simplicity of notations in the proof, we write c instead of c_n . It can be easily shown that

$$E[X] = n(1 - e^{-2c}) + O(1), \quad E[X_i] = E[X_o] = ne^{-c}(1 - e^{-c}) + O(1),$$

and that

$$\text{Var}(X) = O(n), \quad \text{Var}(X_i) = \text{Var}(X_o) = O(n).$$

Introducing $\alpha' = 1 - e^{-2c}$ and $\beta'_i = \beta'_o = e^{-c}(1 - e^{-c})$, we see that

$$X^* := \frac{X - n\alpha'}{n^{1/2}}, \quad X_i^* := \frac{X_i - n\beta'_i}{n^{1/2}}, \quad X_o^* := \frac{X_o - n\beta'_o}{n^{1/2}}$$

a.a.s. are small, i.e. $|X^*|, |X_i^*|, |X_o^*| \leq \ln n$.

The joint distribution of X, X_i, X_o is given by

$$P(X = \nu, X_i = \nu_i, X_o = \nu_o) = \frac{1}{\binom{(n)_2}{m}} \frac{n!}{\nu_i! \nu_o! (\nu - \nu_i - \nu_o)! (n - \nu)!} \sum_{\boldsymbol{\delta}, \boldsymbol{\Delta}} g(\boldsymbol{\delta}, \boldsymbol{\Delta}). \quad (8.1)$$

Here $\boldsymbol{\delta}, \boldsymbol{\Delta}$ denote the generic values of in-degrees and out-degrees of a digraph with m arcs and specified vertex subsets of the respective cardinalities ν_i, ν_o, ν , and $g(\boldsymbol{\delta}, \boldsymbol{\Delta})$ is the total number of such digraphs. In particular,

$$\begin{aligned} \delta_r &> 0, \Delta_r = 0, \quad r \in \{1, \dots, \nu_i\}; \\ \Delta_s &> 0, \delta_s = 0, \quad s \in \{\nu_i + 1, \dots, \nu_i + \nu_o\}; \\ \Delta_t &> 0, \delta_t > 0, \quad t \in \{\nu_i + \nu_o + 1, \dots, \nu\}; \\ \delta_u &= \Delta_u = 0, \quad u \in \{\nu + 1, \dots, n\}, \\ \sum_r \delta_r + \sum_t \delta_t &= \sum_s \Delta_s + \sum_t \Delta_t = m. \end{aligned}$$

We find a sharp asymptotic formula for the sum in (8.1) similarly to the asymptotic formula for $g(\mathbf{s})$, see the proof of Theorem 3.5. From Theorem 3.1, we know that

$$g(\boldsymbol{\delta}, \boldsymbol{\Delta}) = H(\boldsymbol{\delta}, \boldsymbol{\Delta}) m! \prod_r \frac{1}{\delta_r!} \prod_s \frac{1}{\Delta_s!} \prod_t \frac{1}{\delta_t! \Delta_t!},$$

where the “fudge factor” $H(\boldsymbol{\delta}, \boldsymbol{\Delta}) \leq 1$. We need to obtain a sharp asymptotic formula for the sum in (8.1), assuming that

$$x_i := \frac{\nu_i - n\beta'_i}{n^{1/2}} = O(\ln n), \quad x_o := \frac{\nu_o - n\beta'_o}{n^{1/2}} = O(\ln n), \quad x := \frac{\nu - n\alpha'}{n^{1/2}} = O(\ln n).$$

First, consider the sum without the factor $H(\boldsymbol{\delta}, \boldsymbol{\Delta})$. Then

$$m! \sum_{\boldsymbol{\delta}, \boldsymbol{\Delta}} \prod_r \frac{1}{\delta_r!} \prod_s \frac{1}{\Delta_s!} \prod_t \frac{1}{\delta_t! \Delta_t!} = m! [x^m y^m] (e^x - 1)^{\nu - \nu_i} (e^y - 1)^{\nu - \nu_o} \quad (8.2)$$

$$\leq m! \frac{(e^z - 1)^{2\nu - \nu_i - \nu_o}}{z^{2m}}, \quad (8.3)$$

for any $z > 0$. Introduce a truncated Poisson Z , where

$$P(Z = j) = \frac{e^{-c} c^j / j!}{1 - e^{-c}} = \frac{c^j / j!}{e^c - 1}, \quad j \geq 1,$$

so that

$$E[z^Z] = \frac{e^{zc} - 1}{e^c - 1}.$$

In particular,

$$E[Z] = \frac{ce^c}{e^c - 1} = \frac{c}{1 - e^{-c}}.$$

Then

$$\begin{aligned} [z^m](e^z - 1)^a &= \frac{(e^c - 1)^a}{c^m} [z^m] \left(\frac{e^{zc} - 1}{e^c - 1} \right)^a = \frac{(e^c - 1)^a}{c^m} [z^m] (E[z^Z])^a \\ &= \frac{(e^c - 1)^a}{c^m} P\left(\sum_{t \in [a]} Z_t = m\right), \end{aligned}$$

where Z_1, \dots, Z_a are independent copies of Z . We use this formula for

$$\begin{aligned} a_i &= \nu - \nu_i = n(1 - e^{-c}) + (x - x_i)n^{1/2}, \\ a_o &= \nu - \nu_o = n(1 - e^{-c}) + (x - x_o)n^{1/2}, \end{aligned}$$

so that

$$\begin{aligned} m - a_i E[Z] &= cn - [n(1 - e^{-c}) + (x - x_i)n^{1/2}] \frac{c}{1 - e^{-c}} \\ &= -\frac{c}{1 - e^{-c}} (x - x_i)n^{1/2}, \\ m - a_o E[Z] &= cn - [n(1 - e^{-c}) + (x - x_o)n^{1/2}] \frac{c}{1 - e^{-c}} \\ &= -\frac{c}{1 - e^{-c}} (x - x_o)n^{1/2}. \end{aligned}$$

Let's use a shorthand $f \sim g$ if

$$f(x_i, x_o, x) = (1 + o(1))g(x_i, x_o, x), \quad \text{uniformly over } x_i, x_o, x = O(\ln n).$$

By the local central limit theorem (see Durrett [13]),

$$\begin{aligned} P\left(\sum_{t \in [a_i]} Z_t = m\right) &= P\left(\sum_{t \in [a_i]} (Z_t - E[Z]) = m - a_i E[Z]\right) \\ &\sim \frac{1}{\sqrt{2\pi a_i \text{Var}(Z)}} \exp\left[-\frac{(m - a_i E[Z])^2}{2a_i \text{Var}(Z)}\right] \\ &\sim \frac{1}{\sqrt{2\pi n(1 - e^{-c}) \text{Var}(Z)}} \exp\left[-\frac{c^2(x - x_i)^2}{2(1 - e^{-c})^3 \text{Var}(Z)}\right], \end{aligned} \tag{8.4}$$

and likewise

$$P\left(\sum_{t \in [a_o]} Z_t = m\right) \sim \frac{1}{\sqrt{2\pi n(1 - e^{-c})\text{Var}(Z)}} \exp\left[-\frac{c^2(x - x_o)^2}{2(1 - e^{-c})^3\text{Var}(Z)}\right]. \quad (8.5)$$

By (8.4)–(8.5), the equation (8.2) becomes

$$\begin{aligned} m! \sum_{\boldsymbol{\delta}, \boldsymbol{\Delta}} \prod_r \frac{1}{\delta_r!} \prod_s \frac{1}{\Delta_s!} \prod_t \frac{1}{\delta_t! \Delta_t!} &= \frac{m!}{c^{2m}} (e^c - 1)^{2\nu - \nu_i - \nu_o} P\left(\sum_{t \in [a_i]} Z_t = m\right) P\left(\sum_{t \in [a_o]} Z_t = m\right) \\ &\sim \frac{m!}{c^{2m}} \cdot \frac{(e^c - 1)^{2\nu - \nu_i - \nu_o}}{2\pi n(1 - e^{-c})\text{Var}(Z)} \exp\left[-\frac{c^2((x - x_i)^2 + (x - x_o)^2)}{2(1 - e^{-c})^3\text{Var}(Z)}\right]. \end{aligned} \quad (8.6)$$

Recall though that for an asymptotic formula for the sum on the RHS in (8.1) we need to engage the fudge factor $H(\boldsymbol{\delta}, \boldsymbol{\Delta})$. To this end, we notice that the dominant contribution to the RHS of (8.6) comes from $\boldsymbol{\delta}, \boldsymbol{\Delta}$ with

$$\|(\boldsymbol{\delta}, \boldsymbol{\Delta})\| := \max_v (\delta_v + \Delta_v) \leq 2 \ln n.$$

Indeed if $d \geq 2 \ln n$, then $d! \geq (d - \lfloor \ln n \rfloor)! (\ln n)^{\ln n}$. So, similarly to (8.3) (with $z := c$), we obtain that the total contribution of $\boldsymbol{\delta}, \boldsymbol{\Delta}$ with $\|(\boldsymbol{\delta}, \boldsymbol{\Delta})\| \geq 2 \ln n$ to the sum in (8.1) is at most of order

$$nm! (\ln n)^{-\ln n} \frac{(e^c - 1)^{2\nu - \nu_i - \nu_o}}{c^{2m}}. \quad (8.7)$$

For $\|(\boldsymbol{\delta}, \boldsymbol{\Delta})\| < 2 \ln n$, we have McKay's formula

$$\begin{aligned} \ln H(\boldsymbol{\delta}, \boldsymbol{\Delta}) &= -\frac{1}{m} \sum_t \Delta_t \delta_t - \frac{1}{2m^2} \left(\sum_r (\delta_r)_2 + \sum_t (\delta_t)_2 \right) \left(\sum_s (\Delta_s)_2 + \sum_t (\Delta_t)_2 \right) \\ &\quad + O(m^{-1} (\ln n)^4). \end{aligned}$$

Defining $H(\boldsymbol{\delta}, \boldsymbol{\Delta}) > 0$ for $\|(\boldsymbol{\delta}, \boldsymbol{\Delta})\| \geq 2 \ln n$, we write

$$\sum_{\substack{\boldsymbol{\delta}, \boldsymbol{\Delta} \\ \|(\boldsymbol{\delta}, \boldsymbol{\Delta})\| < 2 \ln n}} g(\boldsymbol{\delta}, \boldsymbol{\Delta}) = \frac{m!(e^c - 1)^{a_i + a_o}}{c^{2m}} E[H(\mathbf{Z}^i, \mathbf{Z}^o) 1_{\{|\mathbf{Z}^i| = |\mathbf{Z}^o| = m\}}].$$

Here \mathbf{Z}^i (\mathbf{Z}^o resp.) consists of $a_i = \nu - \nu_i$ (resp. $a_o = \nu - \nu_o$) independent copies of Z . Now, with probability $1 - O(n^{-b})$, for any $b > 0$,

$$\sum_t \Delta_t \delta_t \sim (\nu - \nu_i - \nu_o) E^2[Z],$$

and

$$\sum_r (Z_r^i)_2 + \sum_t (Z_t^i)_2 \sim a_o E[(Z)_2], \quad \sum_s (Z_s^o)_2 + \sum_t (Z_t^o)_2 \sim a_i E[(Z)_2].$$

Since

$$P(|\mathbf{Z}^i| = |\mathbf{Z}^o| = m) = \Theta(n^{-1})$$

is only polynomially small, we obtain that

$$E[H(\mathbf{Z}^i, \mathbf{Z}^o) 1_{\{|\mathbf{Z}^i| = |\mathbf{Z}^o| = m\}}] \sim P(|\mathbf{Z}^i| = |\mathbf{Z}^o| = m) \times \exp \left[-\frac{\nu - \nu_i - \nu_o}{m} E^2[Z] - \frac{a_i a_o}{2m^2} E^2[(Z)_2] \right].$$

Combining this estimate with (8.7) we obtain

$$\begin{aligned} \sum_{\delta, \Delta} g(\delta, \Delta) &\sim \exp \left(-\frac{\nu - \nu_i - \nu_o}{m} E^2[Z] - \frac{(\nu - \nu_i)(\nu - \nu_o)}{2m^2} E^2[(Z)_2] \right) \\ &\times \frac{m!}{c^{2m}} \cdot \frac{(e^c - 1)^{2\nu - \nu_i - \nu_o}}{2\pi n(1 - e^{-c}) \text{Var}(Z)} \exp \left[-\frac{c^2((x - x_i)^2 + (x - x_o)^2)}{2(1 - e^{-c})^3 \text{Var}(Z)} \right]. \end{aligned} \quad (8.8)$$

Note that

$$\frac{(e^c - 1)^{2\nu - \nu_i - \nu_o}}{e^{2m}} = \frac{(e^c - 1)^{(2\alpha' - \beta'_i - \beta'_o)n} (e^c - 1)^{(2x - x_i - x_o)\sqrt{n}}}{e^{2cn}},$$

and

$$(e^c - 1)^{2\alpha' - \beta'_i - \beta'_o} e^{-2c} = (e^c - 1)^{2 - 2e^{-c}} e^{-2c} = e^{-2ce^{-c}} (1 - e^{-c})^{2 - 2e^{-c}}.$$

Turn to the remaining factors in (8.1). Stirling formula yields

$$\frac{m!}{\binom{(n)_2}{m}} \sim 2\pi m e^{-c - c^2/2} \left(\frac{c}{e}\right)^{2m}, \quad (8.9)$$

and

$$\begin{aligned} \nu_i! &= (1 + O(1/n)) \sqrt{2\pi\nu_i} \left(\frac{\nu_i}{e}\right)^{\nu_i} \\ &\sim \sqrt{2\pi\beta'_i n} \frac{1}{e^{\nu_i}} \exp \left((\beta'_i n + x_i \sqrt{n}) \left(\ln n + \ln \beta'_i + \ln \left(1 + \frac{x_i}{\beta'_i \sqrt{n}} \right) \right) \right) \\ &\sim C_i \sqrt{n} \frac{1}{e^{\nu_i}} \exp \left(\beta'_i n \ln n + \beta'_i n \ln \beta'_i + x_i \sqrt{n} \ln n + x_i \sqrt{n} \ln \beta'_i + x_i \sqrt{n} \right. \\ &\quad \left. + x_i^2/\beta'_i - \frac{1}{2} x_i^2/\beta'_i \right), \end{aligned}$$

for some constant $C_i > 0$. There are similar asymptotic formulas for $\nu_o!$, $(\nu - \nu_i - \nu_o)!$ and $(n - \nu)!$. After simple algebra, we find that for some constant $C > 0$,

$$\begin{aligned} \nu_i! \nu_o! (\nu - \nu_i - \nu_o)! (n - \nu)! &\sim \frac{Cn^2}{e^n} \\ &\times \exp \left\{ n \ln n + n \ln \left(e^{-2ce^{-c}} (1 - e^{-c})^{2 - 2e^{-c}} \right) + \sqrt{n} (2x - x_i - x_o) \ln (e^c - 1) \right. \\ &\quad \left. + \frac{1}{2} \left(\frac{x^2}{1 - \alpha'} + \frac{(x - x_i - x_o)^2}{\alpha' - \beta'_i - \beta'_o} + \frac{x_i^2}{\beta'_i} + \frac{x_o^2}{\beta'_o} \right) \right\}. \end{aligned} \quad (8.10)$$

Using (8.8), (8.9) and (8.10), we find that (8.1) becomes, after simplifying,

$$P(X = \nu, X_i = \nu_i, X_o = \nu_o) = (1 + o(1))A(2\pi n)^{-3/2} \exp\left(-\frac{1}{2}\mathbf{x}^T \mathbf{R} \mathbf{x}\right); \quad (8.11)$$

here $\mathbf{x} = (\nu, \nu_i, \nu_o)^T$, $A = A(c) > 0$ and the *symmetric positive-definite* 3×3 matrix \mathbf{R} is defined by

$$\begin{aligned} \mathbf{x}^T \mathbf{R} \mathbf{x} = & \frac{c^2 (x - x_i)^2}{2(1 - e^{-c})^3 \text{Var}[Z]} + \frac{c^2 (x - x_o)^2}{2(1 - e^{-c})^3 \text{Var}[Z]} + \frac{x^2}{1 - \alpha'} \\ & + \frac{(x - x_i - x_o)^2}{\alpha' - \beta'_i - \beta'_o} + \frac{x_i^2}{\beta'_i} + \frac{x_o^2}{\beta'_o}. \end{aligned}$$

That $o(1)$ term is small uniformly over (x, x_i, x_o) with $|x| + |x_i| + |x_o| \leq B \ln n$, $B > 0$ being fixed. So (8.11) is actually the local limit theorem for (X, X_i, X_o) , and so necessarily $A = \sqrt{\det \mathbf{R}}$. Consequently

$$\left(\frac{X - \alpha' n}{\sqrt{n}}, \frac{X_i - \beta'_i n}{\sqrt{n}}, \frac{X_o - \beta'_o n}{\sqrt{n}} \right) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, \mathbf{K}),$$

where $\mathbf{K} = \mathbf{R}^{-1}$. □

Using Mathematica, we found

$$\mathbf{K}(c) = \mathbf{R}^{-1} = \frac{e^c - 1}{e^{4c}(2e^c - c - 2)} \begin{pmatrix} F & G & G \\ G & H & I \\ G & I & H \end{pmatrix},$$

where

$$\begin{aligned} F &= 2e^{2c} - ce^c - 2 - 3c, \\ G &= (2 + c)e^c - 2 - 3c, \\ H &= 2e^{3c} + (-4 - 2c)e^{2c} + (4 + 3c)e^c - 3c - 2, \\ I &= -2e^c + 2 + 3c. \end{aligned}$$

9 Combining Distribution Results

9.1 Asymptotic Distribution of the $(1, 1)$ -core in $D(n, m = c_n n)$

In this Section, we combine the asymptotic distribution of $(\bar{\nu}, \bar{\mu})$ for the process with a generic $\mathbf{s}(0)$ (Section 7) and the asymptotic distribution of $\mathbf{s}(0)$ in $D(n, m = c_n n)$, just proved in Section 8, to show that, for the random $\mathbf{s}(0)$, $(\bar{\nu}, \bar{\mu})$ is asymptotically Gaussian as well (Theorem 1.3). The argument is somewhat similar to the proof of Lemma 2 in Pittel [30].

Theorem 9.1. Let $\bar{\nu}, \bar{\mu}$ be the number of vertices and arcs left at the end of the process starting from $D(n, m = c_n n)$. Denote $\theta_n = \theta(c_n)$. Then we have that

$$\left(\frac{\bar{\nu} - \theta_n^2 n}{\sqrt{n}}, \frac{\bar{\mu} - c_n \theta_n^2 n}{\sqrt{n}} \right) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, \mathbf{B}),$$

where $\mathcal{N}(\mathbf{0}, \mathbf{B})$ is a Gaussian distribution with mean $\mathbf{0}$ and a covariance matrix $\mathbf{B} = \mathbf{B}(c)$ which is continuous in c .

Proof. Let $\mathbf{t} \in \mathbb{R}^2$. We wish to show that

$$\bar{\varphi}(\mathbf{t}) := E \left[\exp \left(it_1 \frac{\bar{\nu} - \theta_n^2 n}{\sqrt{n}} + it_2 \frac{\bar{\mu} - c_n \theta_n^2 n}{\sqrt{n}} \right) \right] = \exp \left(-\frac{1}{2} \mathbf{t}^T \mathbf{B} \mathbf{t} \right) + o(1),$$

where this expectation corresponds to the process starting at random $\mathbf{s}(0)$ from $D(n, m = c_n n)$. Since $D(n, m = c_n n)$ conditioned on $\{\mathbf{s}(0) = \mathbf{s}\}$ is uniform among possible digraphs with parameter \mathbf{s} , we have

$$\bar{\varphi}(\mathbf{t}) = \sum_{\mathbf{s}} E_{\mathbf{s}} \left[\exp \left(it_1 \frac{\bar{\nu} - \theta_n^2 n}{\sqrt{n}} + it_2 \frac{\bar{\mu} - c_n \theta_n^2 n}{\sqrt{n}} \right) \right] P(\mathbf{s}(0) = \mathbf{s}).$$

Let's denote the generic expectation in the sum by $\bar{\varphi}_{\mathbf{s}}(\mathbf{t})$. We break the sum into two parts, for the likely \mathbf{s} and for the unlikely \mathbf{s} . For the first part, we pick $a > 0$ and let $\mathbf{S}(a)$ be the set of $\mathbf{s} = (\nu, \nu_i, \nu_o, \mu = c_n n)$ such that

$$\sqrt{(\nu - \alpha' n)^2 + (\nu_i - \beta'_i n)^2 + (\nu_o - \beta'_o n)^2} \leq a \sqrt{n}, \quad (9.1)$$

where $\alpha' = 1 - e^{-2c_n}$ and $\beta'_i = \beta'_o = e^{-c_n} (1 - e^{-c_n})$. One can easily show that $\mathbf{S}(a) \subset \mathbf{S}_{\epsilon'}$ (defined in (3.9)), where $\epsilon' = n^{-1/3}$. For the second part, we consider the remaining \mathbf{s} , i.e. $\mathbf{s} \notin \mathbf{S}(a)$. For these unlikely \mathbf{s} ,

$$\left| \sum_{\mathbf{s} \notin \mathbf{S}(a)} \bar{\varphi}_{\mathbf{s}}(\mathbf{t}) P(\mathbf{s}(0) = \mathbf{s}) \right| \leq \sum_{\mathbf{s} \notin \mathbf{S}(a)} P(\mathbf{s}(0) = \mathbf{s}) \rightarrow P(|\mathcal{N}(\mathbf{0}, \mathbf{K})| \geq a), \quad (9.2)$$

which is small if a is large.

For $\mathbf{s} \in \mathbf{S}(a)$, we can rewrite $\bar{\varphi}_{\mathbf{s}}(\mathbf{t})$ as

$$\begin{aligned} \bar{\varphi}_{\mathbf{s}}(\mathbf{t}) &= \exp \left(i \mathbf{t}^T (\mathbf{f}(\mathbf{s}/n) - (\theta_n^2, c_n \theta_n^2)^T) \sqrt{n} \right) \\ &\quad \times E_{\mathbf{s}} [\exp (i \mathbf{t}^T ((\bar{\nu}, \bar{\mu})^T - n \mathbf{f}(\mathbf{s}/n)) n^{-1/2})]. \end{aligned} \quad (9.3)$$

In the proof of Corollary 7.3, (see (7.9)), we found that uniformly over $\mathbf{s} \in \mathbf{S}_{\epsilon'}$,

$$E_{\mathbf{s}} [\exp (i \mathbf{t}^T ((\bar{\nu}, \bar{\mu})^T - n \mathbf{f}(\mathbf{s}/n)) n^{-1/2})] = \exp \left(-\frac{1}{2} \mathbf{t}^T \boldsymbol{\psi}(\mathbf{s}/n) \mathbf{t} \right) + O(n^{-0.49}). \quad (9.4)$$

By (9.1), uniformly for $\mathbf{s} \in \mathbf{S}(a)$, we have that $\|\mathbf{s}/n - (\alpha', \beta'_i, \beta'_o, c_n)^T\| = O(n^{-1/2})$, whence

$$\boldsymbol{\psi}(\mathbf{s}/n) = \boldsymbol{\psi}(\alpha', \beta'_i, \beta'_o, c_n) + O(n^{-1/2}), \quad (9.5)$$

because ψ has bounded partial derivatives near $(\alpha', \beta'_i, \beta'_o, c_n)$. Combining (9.4) and (9.5), we rewrite (9.3) as

$$\bar{\varphi}_{\mathbf{s}}(\mathbf{t}) = \exp\left(i\mathbf{t}^T (\mathbf{f}(\mathbf{s}/n) - (\theta_n^2, c_n \theta_n^2)^T) \sqrt{n}\right) \cdot \exp\left(-\frac{1}{2}\mathbf{t}^T \boldsymbol{\psi} \mathbf{t}\right) + O(n^{-.49}), \quad (9.6)$$

where $\boldsymbol{\psi}$ is evaluated at $(\alpha', \beta'_i, \beta'_o, c_n)$.

However the fluctuation of f_j around $(\alpha', \beta'_i, \beta'_o, c_n)$ is on the order of $n^{-1/2}$ and we multiply f_j by \sqrt{n} in (9.6), so we must take this variation into account. If we denote the gradient at f_j at the point $(\alpha', \beta'_i, \beta'_o, c_n)$ by ∇f_j , then the Taylor expansion of f_1 around $(\alpha', \beta'_i, \beta'_o, c)$ gives: uniformly over $\mathbf{s} \in \mathbf{S}(a)$,

$$\sqrt{n} (f_1(\mathbf{s}/n) - \theta_n^2) = \nabla f_1 \cdot \left(\frac{\nu - \alpha' n}{\sqrt{n}}, \frac{\nu_i - \beta'_i n}{\sqrt{n}}, \frac{\nu_o - \beta'_o n}{\sqrt{n}}, 0 \right) + O\left(\frac{1}{\sqrt{n}}\right),$$

and

$$\sqrt{n} (f_2(\mathbf{s}/n) - c_n \theta_n^2) = \nabla f_2 \cdot \left(\frac{\nu - \alpha' n}{\sqrt{n}}, \frac{\nu_i - \beta'_i n}{\sqrt{n}}, \frac{\nu_o - \beta'_o n}{\sqrt{n}}, 0 \right) + O\left(\frac{1}{\sqrt{n}}\right).$$

Hence, uniformly over $\mathbf{s} \in \mathbf{S}(a)$, we have

$$\begin{aligned} \exp\left(i\mathbf{t}^T (\mathbf{f}(\mathbf{s}/n) - (\theta_n^2, c_n \theta_n^2)^T) \sqrt{n}\right) &= \\ \exp\left(iu_1 \frac{\nu - \alpha' n}{\sqrt{n}} + iu_2 \frac{\nu_i - \beta'_i n}{\sqrt{n}} + iu_3 \frac{\nu_o - \beta'_o n}{\sqrt{n}}\right) &+ O(n^{-0.49}), \end{aligned} \quad (9.7)$$

where $u_1 = t_1(f_1)_\alpha + t_2(f_2)_\alpha$, $u_2 = t_1(f_1)_{\beta_i} + t_2(f_2)_{\beta_i}$, $u_3 = t_1(f_1)_{\beta_o} + t_2(f_2)_{\beta_o}$ with these partial derivatives being evaluated at $(\alpha', \beta'_i, \beta'_o, c_n)$. Therefore, by Lemma 8.1, in conjunction with (9.4) and (9.7),

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{\mathbf{s} \in \mathbf{S}(a)} \bar{\varphi}_{\mathbf{s}}(\mathbf{t}) P(\mathbf{s}(0) = \mathbf{s}) &= \\ \exp\left(-\frac{1}{2}\mathbf{t}^T \boldsymbol{\psi} \mathbf{t}\right) \times \int_{|\mathbf{x}| \leq a} \frac{1}{(2\pi)^{3/2} \sqrt{\det \mathbf{K}}} \exp\left(i\mathbf{u} \cdot \mathbf{x} - \frac{1}{2}\mathbf{x}^T \mathbf{R} \mathbf{x}\right) d\mathbf{x}. \end{aligned} \quad (9.8)$$

Furthermore, we notice that

$$\int \frac{1}{(2\pi)^{3/2} \sqrt{\det \mathbf{K}}} \exp\left(i\mathbf{u} \cdot \mathbf{x} - \frac{1}{2}\mathbf{x}^T \mathbf{R} \mathbf{x}\right) d\mathbf{x} = \exp\left(-\frac{1}{2}\mathbf{u}^T \mathbf{K} \mathbf{u}\right).$$

Letting $a \rightarrow \infty$ in (9.2) and (9.8), we obtain that

$$\lim_{n \rightarrow \infty} \bar{\varphi}(\mathbf{t}) = \exp\left(-\frac{1}{2}\mathbf{t}^T \boldsymbol{\psi} \mathbf{t} - \frac{1}{2}\mathbf{u}^T \mathbf{K} \mathbf{u}\right).$$

Equivalently,

$$\lim_{n \rightarrow \infty} \bar{\varphi}(\mathbf{t}) = \exp\left(-\frac{1}{2}\mathbf{t}^T \mathbf{B} \mathbf{t}\right),$$

where $\mathbf{B}(c) = (B_{j,k})$ is the 2×2 matrix defined by

$$\mathbf{B} = \begin{pmatrix} \psi_{1,1} & \psi_{1,2} \\ \psi_{2,1} & \psi_{2,2} \end{pmatrix} + \begin{pmatrix} (f_1)_\alpha & (f_1)_{\beta_i} & (f_1)_{\beta_o} \\ (f_2)_\alpha & (f_2)_{\beta_i} & (f_2)_{\beta_o} \end{pmatrix} \mathbf{K} \begin{pmatrix} (f_1)_\alpha & (f_2)_\alpha \\ (f_1)_{\beta_i} & (f_2)_{\beta_i} \\ (f_1)_{\beta_o} & (f_2)_{\beta_o} \end{pmatrix}, \quad (9.9)$$

and all functions involved are evaluated at

$$\mathbf{w}' := (\alpha', \beta'_i, \beta'_o, c) = (1 - e^{-2c}, e^{-c} - e^{-2c}, e^{-c} - e^{-2c}, c). \quad (9.10)$$

It is straightforward to compute the partial derivatives of f_1 and f_2 from their definition in Proposition 5.7 and we omit the explicit formulas. Clearly $\mathbf{B}(c)$ is continuous since each of its components are. \square

9.2 Description of the mean and covariance parameters

The limiting covariance matrix $\mathbf{B}(c)$, defined in (9.9), of the scaled random vector $(\bar{\nu}, \bar{\mu})$, or in other words $(|V_{1,1}|, |A_{1,1}|)$, is written as a sum of $\boldsymbol{\psi}$, whose entries were given as integral expressions, and a matrix whose entries are explicit functions of c (in the next section, we will prove that $(|V_1|, |A_1|)$ has the same asymptotic distribution). We defined the entries $\psi_{j,k}$ using the method of characteristics and found that

$$\psi_{j,k}(\mathbf{w}_0) = \int_{z^*(0)}^{z_i(0)} \mathcal{E}_s[(\Delta^T \nabla f_j)(\Delta^T \nabla f_k)] \frac{(\beta_i + \beta_o)(\alpha - \beta_o)}{\beta_i z_i} dz_i. \quad (9.11)$$

For general \mathbf{w}_0 , since each of $\alpha, \beta_i, \beta_o, \gamma, z_i, z_o$ are strictly decreasing along the trajectory, each variable can be written as a function of z_i . However, determining these exact expressions is difficult to do for general \mathbf{w}_0 . Fortunately, for \mathbf{w}' , defined in (9.10), we can find these expressions.

Since the trajectory starting at \mathbf{w}' begins with $\beta_i(0) = \beta_o(0)$, by Proposition 5.4, we have that $\beta_i(t) = \beta_o(t) =: \beta(t)$ for $t \geq 0$; whence $z_i(t) = z_o(t) =: z(t)$. Further, we found that the functions I_1 and I_2 are constants along this trajectory; in particular,

$$I_1(\mathbf{w}') = c = \frac{\gamma(\alpha - 2\beta)}{(\alpha - \beta)^2}, \quad I_2(\mathbf{w}') = c = \frac{z^2}{\gamma}.$$

These two equations along with the definition of z (i.e. $\frac{ze^z}{e^z - 1} = \frac{\gamma}{\alpha - \beta}$) allow us to solve for α, β , and γ in terms of z . Thus the integrand in (9.11) can be written as an explicit function of z over the interval $[z^*(\mathbf{w}'), z_i(\mathbf{w}')] = [c\theta, c]$. However, even these simplified versions of the integrals are rather long and are omitted. Nevertheless, one can approximate the entries of $\boldsymbol{\psi}$ by using numerical integration. We refer the interested reader to a Mathematica notebook file, which contains these integrands at <http://www.dpoole.info/strong-giant/>.

We can say much more about $\psi_{j,k}(\mathbf{w}')$ for $c = 1 + \epsilon$, $\epsilon \downarrow 0$. Using Mathematica, we expand the integrand around $z = 0$ and $\epsilon = 0$. This allows us to integrate this power series for $z \in [c\theta, c]$ and determine the leading order terms. We find that $\psi_{j,k} =$

$40\epsilon + O(\epsilon^2)$, for $j, k \in 1, 2$. So while the individual variances of $|V_1|$ and $|A_1|$ are both of order $n\epsilon$, the variance of the excess $\text{Exc}_1 := |A_1| - |V_1|$ is of order $n\epsilon^2$, at most. So, instead of $(|V_1|, |A_1|)$, we consider the random vector $(|V_1|, \text{Exc}_1)$. This random vector is asymptotically Gaussian with mean $n(\theta^2, (c-1)\theta^2)$ and covariance matrix $n\tilde{\mathbf{B}}$, where

$$\tilde{\mathbf{B}} := \begin{pmatrix} B_{1,1} & B_{1,2} - B_{1,1} \\ B_{1,2} - B_{1,1} & B_{1,1} + B_{2,2} - 2B_{1,2} \end{pmatrix}. \quad (9.12)$$

As $\epsilon \downarrow 0$, the leading order terms for the mean vector are $n(4\epsilon^2, 4\epsilon^3)$. To obtain the leading order terms for the covariance matrix, we needed to determine $\psi_{j,k}$ up to order ϵ^3 . In particular, we find that

$$\tilde{\mathbf{B}} = \begin{pmatrix} 40\epsilon + O(\epsilon^2) & 60\epsilon^2 + O(\epsilon^3) \\ 60\epsilon^2 + O(\epsilon^3) & \frac{272}{3}\epsilon^3 + O(\epsilon^4) \end{pmatrix}. \quad (9.13)$$

Thus the variance of Exc_1 is actually of order $n\epsilon^3$, rather than naively expected $n\epsilon^2$. The leading order terms for this covariance matrix $\tilde{\mathbf{B}}$ come from the contributions of $\boldsymbol{\psi}$ rather than those of the second term of (9.9). The latter ones are felt only in the next higher order terms.

Qualitatively, the parameters of the largest strong component in $D(n, m = cn)$ are closer to the parameters of the giant 2-core in $G(n, m = cn/2)$, than to those of $G(n, m = cn/2)$'s giant component. Indeed, from the results in [36] it follows that, for $c = 1 + \epsilon$, $\epsilon \downarrow 0$, the scaled mean vector and covariance matrix of the (asymptotically Gaussian) pair (size of 2-core, excess of 2-core) for $G(n, m = cn/2)$ are of exactly the same orders as their counterparts for $(|V_1|, \text{Exc}_1)$ in $D(n, m = cn)$.

9.3 Asymptotic Distribution of the $(1, 1)$ -core in $D(n, p = c_n/n)$

As essentially an afterthought, here is a similar claim for $D(n, p = c_n/n)$.

Theorem 9.2. *Let $V_{1,1}^p, A_{1,1}^p$ denote the vertex set and arc set of the $(1, 1)$ -core in the random digraph $D(n, p)$. Suppose $p = c_n/n$ and $\lim c_n = c \in (1, \infty)$. Denote $\boldsymbol{\mu}(c) = (\theta^2(c), c\theta^2(c))$. Note that $\mu(x)$ is differentiable and we denote its derivative as $\boldsymbol{\mu}'(x)$. Then*

$$\left(\frac{|V_{1,1}^p| - \theta_n^2 n}{n^{1/2}}, \frac{|A_{1,1}^p| - c_n \theta_n^2 n}{n^{1/2}} \right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathcal{B}),$$

where $\mathcal{N}(\mathbf{0}, \mathcal{B})$ is the 2-dimensional Gaussian vector with mean $\mathbf{0}$ and the 2×2 covariance matrix \mathcal{B} , given by

$$\mathcal{B}(c) = \mathbf{B}(c) + c(\boldsymbol{\mu}'(c))^T(\boldsymbol{\mu}'(c)).$$

The proof is basically a copy of the argument in [30] (Lemma 2), that allows transfer of the asymptotic normality results for the Erdős-Rényi graph $G(n, m)$ to the Bernoulli model $G(n, p)$, with $p = m/\binom{n}{2}$. For the sake of completeness, we include the proof in the Appendix (see Lemma A.1).

Consequently, the pair $(|V_1|, \text{Exc}_1)$ is asymptotically Gaussian with mean $n(\theta^2), (c - 1)\theta^2$) and the covariance matrix $\mathcal{B}'(c)$ obtained from $\mathcal{B}(c)$ just like \mathbf{B}' is obtained from \mathbf{B} , see (9.12). For $\epsilon = c - 1 \downarrow 0$, \mathcal{B}' is given by the asymptotic expression on the RHS of (9.13).

Thus we have proved that for both $D(n, m = c_n n)$ and $D(n, p = c_n/n)$, the pair $(|V_{1,1}|, |A_{1,1}|)$ is asymptotically Gaussian, with the covariance matrix dependent on the model. What's left is to show that in each model a.a.s. the parameters $(|V_{1,1}|, |A_{1,1}|)$ of the $(1, 1)$ -core are at a polylog distance from the parameters $(|V_1|, |A_1|)$ of the strong giant, the distance negligible compared to the likely random fluctuations of the $(1, 1)$ -core's parameters. Our argument will utilize a depth-first-search process, rather than the deletion process we have used for the $(1, 1)$ -core.

10 Structure of $(1, 1)$ -core in $D(n, m)$ and $D(n, p)$

The main goal of this section is to prove that a.a.s. the $(1, 1)$ -core of $D(n, m = c_n n)$ and of $D(n, p = c_n/n)$, contains at most a polylog number of vertices and arcs outside the largest strong component (Theorem 1.2). Breaking with the logic sequence so far, we first prove this theorem for $D(n, p)$ and then transfer this result to $D(n, m)$.

Let's define $\mathcal{D}(V)$ ($\mathcal{A}(V)$ resp.) to be the descendant (ancestor resp.) of a set of vertices V . As a reminder, we define a vertex to be a descendant (ancestor) of itself. So by definition, for any vertex set V , we have that $V \subset \mathcal{D}(V) \cap \mathcal{A}(V)$; further if V is the vertex set of a strong component, then $V = \mathcal{D}(V) \cap \mathcal{A}(V)$.

To prove Theorem 1.2, it is sufficient to prove the following lemma. Let us say that a event A holds *asymptotically more surely* (a.m.s.) if $P(A) \geq 1 - o(n^{-1/2})$. We recall that an event A holds q.s., if $P(A) \geq 1 - O(n^{-a})$ for any fixed $a > 0$.

Lemma 10.1. *A.m.s., $D(n, p = c_n/n)$ is such that*

$$|V_{1,1} \setminus \mathcal{D}(V_1)| \leq (\ln n)^8, \quad |V_{1,1} \setminus \mathcal{A}(V_1)| \leq (\ln n)^8. \quad (10.1)$$

Before proving this lemma, let us show how it implies Theorem 1.2.

Proof of Theorem 1.2. Since $V_1 = \mathcal{D}(V_1) \cap \mathcal{A}(V_1)$, we have that

$$V_{1,1} \setminus V_1 = (V_{1,1} \setminus \mathcal{D}(V_1)) \cup (V_{1,1} \setminus \mathcal{A}(V_1)),$$

and so, a.m.s.,

$$|V_{1,1}| - |V_1| = |V_{1,1} \setminus V_1| \leq 2(\ln n)^8. \quad (10.2)$$

Furthermore, one can easily show that q.s. max in-degree and max out-degree is at most $(\ln n)^2$; on this event, each vertex in $V_{1,1} \setminus V_1$ is an endpoint of at most $(\ln n)^2 + (\ln n)^2$ arcs. Hence, a.m.s.,

$$0 \leq |A_{1,1}| - |A_1| \leq 2(\ln n)^8 \cdot 2(\ln n)^2 = 4(\ln n)^{10}. \quad (10.3)$$

A standard inequality

$$P(D(n, m) \in A) = O \left[\sqrt{\text{Var}(e(D(n, p)))} P(D(n, p) \in A) \right],$$

for any event A , where $e(D)$ is the number of arcs of D and $p = m/n^2$, and $\text{Var}(e(D(n, p = c_n/n))) = O(n^{1/2})$, allow us to convert the a.m.s. bounds in (10.2), (10.3) for $D(n, p = c_n/n)$ into a.a.s. bounds in $D(n, m = c_n n)$. \square

Let us outline the proof of Lemma 10.1. By symmetry, it suffices to prove only the first bound in (10.1). Following T. Łuczak and Seierstad [24], we partition the vertex set of $D(n, p)$ into 2 sets, V and $[n] \setminus V$, where V is the set of all vertices that are descendants of “long” cycles, i.e. those of length at least $(\ln n)^2$. All other cycles are called short.

Vertices of $[n] \setminus V$ can not be reached from V , because otherwise these vertices would be descendants of long cycles. Furthermore, the induced digraph on $[n] \setminus V$ can not contain a long cycle.

Recall that a vertex is in the $(1, 1)$ -core if and only if it is in a cycle or in a directed path joining two cycles. Since there are no arcs from V to $[n] \setminus V$, any vertex in $V_{1,1} \setminus V$ must be a descendant of a necessarily short cycle, with all its vertices in $[n] \setminus V$. So to prove that a.m.s. $|V_{1,1} \setminus V|$ is small, it suffices to show that a.m.s. the descendant set of cycles in the induced digraph on $[n] \setminus V$ is small.

Let \mathcal{L} be the induced digraph on V . The random digraph $D(n, p)$, conditioned on \mathcal{L} , is described as follows:

- there are no arcs from V to $[n] \setminus V$,
- each of the potential arcs from $[n] \setminus V$ to V is present with probability p independently of all other such arcs, as there is no new information on these arcs gained from knowing the set V ,
- the arc set within $[n] \setminus V$ is distributed as $D(n - |V|, p)$, conditioned on the event that there are no long cycles.

Here is our plan for action. First, we show that q.s. $|V| = \theta_n n + O(n^{1/2} \ln n)$. Second, we use this result to prove that q.s. $D(n - |V|, p)$ does not have a long cycle; this allows us to drop conditioning on the event “no long cycles” in the Łuczak-Seierstad decomposition. Third, we show that a.m.s. the descendant set of cycles within $D(n - |V|, p)$ has cardinality at most $(\ln n)^8$. Finally, we prove that a.m.s. $V = \mathcal{D}(V_1)$.

Step 1. We will use the following fundamental result due to Karp [18]. Let $d(v)$ and $c(v)$ denote, respectively, the number of descendants of vertex v in $D(n, p)$ and the size of the component containing v in $G(n, p)$. Then $d(v)$ and $c(v)$ are equidistributed. As for $\{c(v)\}_{v \in [n]}$, the following theorem was proved in [30]:

Theorem 10.2 (Gap Theorem in $G(n, p)$). *Suppose $c'_n \in (1, \infty)$ is bounded away from 1 and ∞ . Then, q.s. exactly one component of $G(n, p = c'_n/n)$ has size in*

$$[\theta'_n n - n^{1/2} \ln n, \theta'_n n + n^{1/2} \ln n],$$

such that $\theta'_n = \theta(c'_n)$, where $\theta(x)$ is the unique root of $1 - \theta = e^{-x\theta}$, and each other component has size less than $(\ln n)^2$.

This theorem and Karp's identity allows us to obtain a similar claim for the sequence of descendant sets delivered by a “full” directed version of the classical depth-first search (DFS) algorithm on $D(n, p)$.

Here is how this algorithm works. Given a directed graph D on $[n]$, we start with vertex $w_1 = 1$ and run the directed depth-first search, finding the chronologically ordered sequence w_2, \dots, w_{k_1} , of all descendants of w_1 together with the resulting search tree T_1 . Moreover, we also find W_1, W_2, \dots, W_{k_1} , their respective descendant sets within the search tree T_1 . Next, pick w_{k_1+1} , the lowest-index vertex from $[n] \setminus V(T_1)$, and repeat the above step on $n \setminus V(T_1)$ obtaining T_2 and the descendant sets of vertices within T_2 . And so on. We end up with a sequence of disjoint search trees T_1, T_2, \dots , and the n -long sequence of vertices w_1, \dots, w_n , together with their *partial* descendant sets W_1, \dots, W_n . Note that the *full* descendant set of the tree T_i in the *digraph* is contained within $T_1 \cup \dots \cup T_i$. Since every two vertices of a strong component are descendants of one another, the vertices of a strong component must reside in the same search tree. Applied to an undirected graph G on $[n]$, the DFS delivers as many search trees T_1, T_2, \dots as there are components of G , each T_i spanning its own component.

Karp's observation can be extended to this full search. Namely, the two sequences of search trees (T_1, T_2, \dots) , one for $D(n, p)$ and another for $G(n, p)$, are equidistributed. So, letting $U_i = V(T_i)$ in $D(n, p)$, without any additional effort, we have

Lemma 10.3. *For $D(n, p = c_n/n)$, q.s. exactly one of U_1, U_2, \dots , is giant, i.e. has size in*

$$[\theta_n n - n^{1/2} \ln n, \theta_n n + n^{1/2} \ln n], \quad (10.4)$$

and all other U_j have size less than $(\ln n)^2$.

Since vertices in a strong component must reside in the same U_i , this gap theorem implies that q.s. any strong component with size at least $(\ln n)^2$ must be contained in the unique giant U_i , of size in the interval (10.4). By the definition of long cycles, each long cycle must reside within this giant U_i . On the q.s. event “there is a unique giant U_i ”, let i_0 denote its index. Since no descendants of long cycles can be found in $\cup_{i > i_0} U_i$, by the definition of V , we have that $V \subset \cup_{i \leq i_0} U_i$, and so

$$|V| \leq \sum_{i=1}^{i_0} |U_i| \leq \theta_n n + n^{1/2} \ln n + i_0 (\ln n)^2.$$

Lemma 10.4. *Q.s. $i_0 \leq (\ln n)^2$. Consequently, q.s. $|V| \leq \theta_n n + 2n^{1/2} \ln n$.*

Proof. First of all, for each $k \in [n]$, we have $\{1, 2, \dots, k\} \subset U_1 \cup \dots \cup U_k$. By Karp's extended correspondence, it suffices to show that q.s. one of the first $(\ln n)^2$ vertices of $[n]$ is in the giant component of $G(n, p)$. To this end, let \mathcal{V} denote the vertex set of largest

component of $G(n, p)$. By symmetry, the set \mathcal{V} conditioned on the value of $|\mathcal{V}|$ is uniformly distributed on all subsets of $[n]$ of size $|\mathcal{V}|$. Therefore,

$$\begin{aligned} P(\mathcal{V} \cap \{1, \dots, (\ln n)^2\} = \emptyset \mid |\mathcal{V}| = \ell) &= P(\mathcal{V} \subset \{(\ln n)^2 + 1, \dots, n\} \mid |\mathcal{V}| = \ell) \\ &= \binom{n - (\ln n)^2}{\ell} / \binom{n}{\ell} \leq \exp(-\ell(\ln n)^2/n). \end{aligned}$$

By the gap theorem in $G(n, p)$, q.s., $|\mathcal{V}| = \theta_n n + O(n^{1/2} \ln n)$, and for $\ell = \theta_n n + O(n^{1/2} \ln n)$, the above conditional probability is $O(n^{-a})$ for any fixed $a > 0$. \square

It remains to prove a matching lower bound for $|V|$. At this moment, we don't even know whether V , the descendant set of long cycles, is empty or not. We will prove likely existence of a special long cycle such that its descendant set has cardinality close to $\theta_n n$. As a first step, we will prove likely existence of a special long path, which will later be shown to be contained a special long cycle; this is where our choice of using a *depth*-first search will play a vital role. In particular, we will prove that q.s. the giant descendant tree T_{i_0} has the following “broom-like” structure, an example of which is shown in Figure 1: (1) there is a long path in T_{i_0} , whose “splinters” have at most polylog total size; (2) the descendant set of the endpoint of this path (the “bristles”) is giant; formal statement of this property is in Lemma 10.6. Once this is proved, we will show that, in $D(n, p)$, there is likely an arc from at least one of the bristles to the top end of the handle, resulting in a long cycle whose descendant set includes all the bristles.

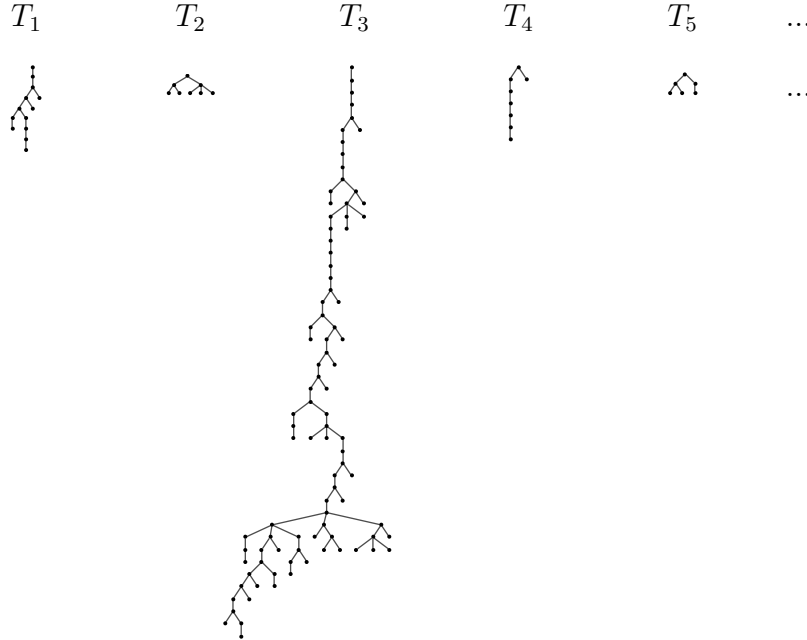


Figure 1: The broom within T_3 in the depth-first search.

First,

Lemma 10.5. *Q.s., for each $i \leq n^{1/2}$, either $|W_i| < (\ln n)^2$ or*

$$|W_i| \in [\theta_n n - 2n^{1/2} \ln n, \theta_n n + 2n^{1/2} \ln n]. \quad (10.5)$$

Proof. By Karp's correspondence, conditioned on the prehistory up to defining w_j , $|W_j|$ is equal in distribution to $c(v)$, the size of the component of a generic vertex v in $G(n - j + 1, p)$. By the Gap Theorem, for each $j \leq n^{1/2}$, q.s. either $|W_j| < (\ln n - j + 1)^2 \leq (\ln n)^2$ or

$$|W_j| \in [\theta_n^{(j)}(n - j + 1) - n^{1/2} \ln n, \theta_n^{(j)}(n - j + 1) + n^{1/2} \ln n], \quad (10.6)$$

where $\theta_n^{(j)} = \theta(c_n(1 - \frac{j-1}{n}))$. For $j \leq n^{1/2}$, we have that $\theta_n^{(j)} = \theta(c_n) + O(n^{-1/2})$ (because the derivative of $\theta(x)$ is bounded near c), which implies that the interval in (10.6) is contained in the interval in (10.5). \square

Next,

Lemma 10.6. (*"broom property"*) *Let $W(v)$ denote the descendant set of a vertex v in the search forest $\{T_i\}$; so $W(w_j) = W_j$. Q.s. in the giant tree T_{i_0} there is a unique path, $y_1 \rightarrow y_2 \rightarrow \dots$ such that*

- (i) *the path starts at the progenitor of T_{i_0} , i.e. $y_1 = \text{root}(T_{i_0})$;*
- (ii) *the path has length $\lceil (\ln n)^3 \rceil$;*
- (iii) *the descendant set of $y_{\lceil (\ln n)^3 \rceil}$ is giant, i.e.*

$$|W(y_{\lceil (\ln n)^3 \rceil})| \in [\theta_n n - 2n^{1/2} \ln n, \theta_n n + 2n^{1/2} \ln n].$$

Proof. Call a vertex v *giant* (*small* resp.) if $|W(v)| \in [\theta_n n - 2n^{1/2} \ln n, \theta_n n + 2n^{1/2} \ln n]$, (if $|W(v)| < (\ln n)^2$, resp.). By the definition of T_{i_0} , y_1 is certainly giant. A giant vertex v can have at most one giant child, because

$$|W(v)| = 1 + \sum_{u: \text{children of } v} |W(u)|.$$

Recursively, if y_j is giant, then let y_{j+1} be the unique giant child if it exists. Let's show that q.s. y_j exists for all $j \leq (\ln n)^3$, and in particular, $y_j = w_k$ for some $k = k(j) \leq j(\ln n)^4$.

Let's consider the event A such that: (a) $i_0 \leq (\ln n)^2$, (b) for each $j \leq n^{1/2}$, w_j is either small or giant, (c) maximum out-degree is less than $(\ln n)^2$. By Lemmas 10.4, 10.5, as well as earlier comments about the event (c), A holds q.s. since every one of the events (a), (b), (c) is a q.s. event.

Suppose the event A holds. Then, by (a), i_0 is defined and $y_1 = w_k$ for $k = |U_1| + \dots + |U_{i_0-1}| + 1$. Furthermore, each of U_1, \dots, U_{i_0-1} have size less than $(\ln n)^2$, so

$$k \leq (i_0 - 1) \cdot (\ln n)^2 + 1 \leq (\ln n)^4.$$

Inductively, suppose that for any $j \in [1, (\ln n)^3]$ we have $y_j = w_k$ with $k \leq j(\ln n)^4$. By (c), y_j has at most $(\ln n)^2$ children. Since y_j is giant, at least one of y_j 's children is not small. Let w_ℓ be the chronologically first child of y_j that is not small, i.e. $|W_\ell| \geq (\ln n)^2 + 1$. Since

we are performing a *depth*-first search, for each earlier, necessary small, child of y_j , we uncover their descendant sets within the search tree before coming to w_ℓ . In particular,

$$\begin{aligned}\ell &= k + 1 + \sum_{v \in \{y'_j \text{'s children before } w_\ell\}} |W(v)| \\ &< k + 1 + ((\ln n)^2 - 1)(\ln n)^2 \leq (j + 1)(\ln n)^4 \leq (\ln n)^8 \leq n^{1/2}.\end{aligned}$$

Then, by (b), a non-small w_ℓ is either small or giant, hence w_ℓ must be giant. \square

Next,

Lemma 10.7. *In $D(n, p = c_n/n)$, q.s. there is an arc from $W(y_{(\ln n)^3})$ to $\{y_1, y_2, \dots, y_{(\ln n)^2}\}$, and $|V| \geq \theta_n n - 2n^{1/2} \ln n$.*

Proof. Conditioned on the search forest $\vec{T} := \{T_i\}$, each possible backward arc, that is an arc (w_j, w_i) for $i < j$, is present in $D(n, p)$ independently with probability p . Let's condition on the event that \vec{T} has the unique path described in Lemma 10.6. By property (iii) of the long path, $|W(y_{(\ln n)^3})| \geq \theta_n n - 2n^{1/2} \ln n$, and so

$$P(\text{no arcs from } W(y_{2(\ln n)^3}) \text{ to } \{y_1, \dots, y_{(\ln n)^2}\} | \vec{T}) = (1 - p)^{|W(y_{(\ln n)^3})|(\ln n)^2} \ll n^{-a},$$

for any fixed $a > 0$. If at least one of these potential arcs is present, then there is a cycle of length at least $(\ln n)^3 - (\ln n)^2 \gg (\ln n)^2$, passing through $y_{(\ln n)^3}$; whence, V is nonempty and $W(y_{(\ln n)^3}) \subset V$. \square

Combining Lemmas 10.4, 10.7, we establish the concentration property of $|V|$.

Corollary 10.8. *Quite surely, $D(n, p = c_n/n)$ is such that*

$$|V| \in [\theta_n n - 2n^{1/2} \ln n, \theta_n n + 2n^{1/2} \ln n],$$

where, we recall, θ_n is the unique positive root of $1 - \theta = e^{-c_n \theta}$.

Step 2.

Lemma 10.9. *Quite surely, there are no long cycles in $D(n - |V|, p = c_n/n)$. Consequently, the probability of any event in $D(n - |V|, p)$, conditioned on having no long cycles, is within distance $O(n^{-a})$ from the probability of this event in the unconditioned $D(n - |V|, p)$.*

Proof. Note that

$$P(\exists \text{ long cycle in } D(n - |V|, p)) = \sum_{\nu=0}^n P(\exists \text{ long cycle in } D(n - \nu, p)) P(|V| = \nu).$$

By Corollary 10.8, q.s. $|V| \in [\theta_n n - 2n^{1/2} \ln n, \theta_n n + 2n^{1/2} \ln n]$. Further, the probability that $D(n - \nu, p)$ contains a long cycle is decreasing in ν , and for $\nu = \nu' := \theta_n n - 2n^{1/2} \ln n$ we have

$$P(\exists \text{ long cycle in } D(n - \nu', p)) \leq \sum_{\ell=(\ln n)^2}^{n-\nu'} \binom{n-\nu'}{\ell} (\ell-1)! p^\ell \leq \sum_{\ell=(\ln n)^2}^{\infty} \frac{1}{\ell} ((n-\nu')p)^\ell.$$

Now $(n - \nu')p = (1 - \theta(c))c + o(1)$ since $c_n \rightarrow c$ and $\theta(x)$ is continuous (see definition of $\theta(x)$ in Theorem 10.2). It is easy to show that $(1 - \theta(c))c < 1$; hence there is some $\delta \in (0, 1)$ such that $(n - \nu')p < \delta$ for all sufficiently large n . Hence

$$P(\exists \text{ long cycle in } D(n - \nu', p)) \leq \sum_{\ell=(\ln n)^2}^{\infty} \frac{\delta^\ell}{\ell} \ll n^{-a}, \quad \forall (\text{fixed}) a > 0.$$

□

Step 3. So we need to bound the size of the descendant set of cycles within $[n] \setminus V$. Let C_n be the number of vertices in short cycles in $D(n - |V|, p)$.

Lemma 10.10. *A.m.s. in $D(n - |V|, p = c_n/n)$, cycles are vertex-disjoint and $C_n \leq (\ln n)^6$.*

Proof. As in the proof of Lemma 10.9, it suffices to consider this event in $D(n', p' = \delta/n')$, where $n' = n - \nu'$, $\nu' = \theta_n n - 2n^{1/2} \ln n$ and $n'p \leq \delta < 1$.

Using the first moment argument, it is straightforward to show that in $D(n', p')$, a.m.s. the number of vertices in *disjoint* short cycles is at most $(\ln n)^6$. Therefore, all that remains is to show that a.m.s. any two cycles are vertex-disjoint. As noticed by Łuczak and Seierstad [24], if two cycles intersect, then there is a directed cycle with a directed chord of possibly zero length. On k vertices, there are at most $k!k^2$ pairs of such cycles and chords. That is, if k_1 denotes the length of the cycle and $k_2 := k - k_1$ denotes the number of vertices in the chord, then the number of such pairs is at most $\binom{k}{k_1} (k_1 - 1)! k_1^2 k_2!$. Summing over k_1 , we find that

$$\sum_{k_1=3}^k \binom{k}{k_1} (k_1 - 1)! k_1^2 k_2! \leq \sum_{k_1=3}^k k! k_1 \leq k! k^2.$$

Therefore, the probability that two cycles intersect in $D(n', p')$ is at most

$$\sum_{k=3}^{n'} \binom{n'}{k} k! k^2 p^{k+1} \leq p \sum_{k=3}^{n'} k^2 \delta^k = O(1/n).$$

□

Having bound the likely number of vertices in cycles in $D(n - |V|, p)$, we turn to the descendant sets of those vertices and complete Step 3.

Lemma 10.11. *In $D(n - |V|, p)$, q.s. for any set of vertices W we have $d(W) \leq |W|(\ln n)^2$. Consequently, by applying Lemmas 10.9, 10.10, in $D(n - |V|, p = c_n/n)$, a.m.s. the size of the descendant set of cycles is at most $(\ln n)^8$. Therefore, in $D(n, p = c_n/n)$, a.m.s.*

$$|V_{1,1} \setminus V| \leq (\ln n)^8.$$

Proof. Again it suffices to consider only $\nu' := \theta_n n - 2n^{1/2} \ln n$. Using Karp's correspondence, and the fact that increasing the edge probability can only increase, stochastically, $c(v)$,

$$P(|d(v)| \geq (\ln n)^2) \leq P(c(v) \geq (\ln n)^2 \text{ in } G(n', p' = \delta/n')),$$

where $n' = n - \nu'$ and $n'p \leq \delta < 1$. Using a first moment argument on the counts of trees of given size, one can show that this latter probability is at most $e^{-b(\ln n')^2} (\ll n^{-a})$ for some $b > 0$. The union bound over all vertices completes the proof. \square

Step 4. Recall that in this Step, we wish to prove that a.m.s., $\mathcal{D}(V_1) = V$. Once proven, it delivers, in combination with Lemma 10.11, the proof of Lemma 10.1.

Lemma 10.12. *In $D(n, p = c_n/n)$, a.m.s. $\mathcal{D}(V_1) = V$.*

Proof. Let V' be the ancestor set of long cycles. The results in Steps (1)-(3) have their obvious analogues to V' . Since there are no arcs from V to $[n] \setminus V$ or from $[n] \setminus V'$ to V' , any strong component outside $V \cap V'$ must completely reside in either $[n] \setminus V$ or $[n] \setminus V'$. Now by Lemma 10.10, a.m.s. each strong component within $[n] \setminus V$ (or $[n] \setminus V'$) is either a single vertex or a short induced cycle, which by the definition of short cycle has less than $(\ln n)^2$ vertices. Therefore, a.m.s. each strong component outside $V \cap V'$ has size less than $(\ln n)^2$. Further by Corollary 10.8, q. s. long cycles do exist, they are necessarily contained in a strong component of size at least $(\ln n)^2$. Hence a.m.s. the largest strong component is within $V \cap V'$. To complete the proof, it suffices to show that a.m.s. $V \cap V'$ is in fact the vertex set of a strong component. To this end, we want to show that a.m.s. between any two long cycles C_1 and C_2 , there are directed paths from C_1 to C_2 and from C_2 to C_1 .

Using Karp's correspondence, by the gap theorem, as well as the union bound over all vertices in $D(n, p)$, we have that q.s., for every vertex v , the descendant set $d(v)$ and the ascendant set $a(v)$ have their cardinalities

$$|d(v)|, |a(v)| \in [1, (\ln n)^2] \cup [\theta_n n - n^{1/2} \ln n, \theta_n n + n^{1/2} \ln n]. \quad (10.7)$$

Given two vertices $v \neq w$, $A(v, w)$ be the event that v and w are both in long cycles and satisfy (10.7), but there is no path from w to v . $A(v, w)$ is contained in the event that $d(v), a(w) = \theta_n n + O(n^{1/2} \ln n)$ but there is no path from v to w . Note that in $D(n, p)$, conditioned on $\mathcal{D}(v)$, the induced subgraph on $[n] \setminus \mathcal{D}(v)$ is distributed as $D(n - |\mathcal{D}(v)|, p)$, without any conditioning. Therefore

$$P(A|\mathcal{D}(v)) \leq P(a(w) = \theta_n n + O(n^{1/2} \ln n) \text{ in } D(n - |\mathcal{D}(v)|, p)).$$

Just as we proved Lemma 10.11, one can easily prove that for $d(v) = \theta_n n + O(n^{1/2} \ln n)$, this latter probability is less than n^{-a} for any fixed $a > 0$. The union bound over all pairs of vertices completes this proof. \square

Acknowledgement. We express our heartfelt gratitude to Huseyin Acan, Nick Peterson and Chris Ross for their invaluable feedback during the last years, and to Neil Falkner and to Matt Kahle, the members of the second author dissertation committee, for their incredible patience and unflagging support. It is a pleasure to thank Nick for his generous help with Mathematica and graphics. We thank the referees for their time and effort required to read this paper with such a painstaking care and for providing us with many thoughtful critical remarks and helpful suggestions.

References

- [1] J. Aronson, A. Frieze, B. Pittel. Maximum matchings in sparse graphs: Karp-Sipser re-visited, *Random Structures & Algorithms*, **12**, (1998) 111-177.
- [2] D. Barraez, S. Boucheron, W. Fernandez De La Vega. On the fluctuations of the giant component, *Combinatorics Probability and Computing*, **9** (2000), 287-304.
- [3] M. Behrisch, A. Coja-Oghlan, M. Kang. The order of the giant component of random hyper graphs, *Random Structures & Algorithms*, **36** (2010), 149-184.
- [4] M. Behrisch, A. Coja-Oghlan, M. Kang. Local limit theorems and the number of connected hypergraphs, *Preprint available from arXiv:0706.0496*, (2007).
- [5] E. A. Bender, E. R. Canfield, B. D. McKay. The asymptotic number of labeled connected graphs with a given number of vertices and edges, *Random Structures & Algorithms*, **1** (1990), 127-169.
- [6] M. Biskup, L. Chayes, S.A. Smith. Large-deviations, thermodynamic approach to percolation on the complete graph, *Random Structures & Algorithms*, **31** (2007), 354-370.
- [7] B. Bollobás. The evolution of random graphs. *Trans. Amer. Math. Soc.* **286** (1984), 257-274.
- [8] B. Bollobás. The evolution of sparse graphs, In *Graph Theory and Combinatorics*. Proc. Cambridge Conf. in honor of Paul Erdős (Bollobás, B., Ed.) Academic Press, pp. 35-37.
- [9] B. Bollobás, O. Riordan. Asymptotic normality of the size of the giant component via a random walk, *J. Combin. Theory Series B*, **102** (2012), 53-61.
- [10] B. Bollobás, O. Riordan. Asymptotic normality of the size of the giant component in a random hypergraph, *Random Structures & Algorithms*, **41** (2012), 441-450.
- [11] C. Cooper, A. Frieze. The size of the largest strongly connected component of a random digraph with a given degree sequence, *Combinatorics, Probability & Computing*, **13** (2004), 319-337.

- [12] J. Ding, J. H. Kim, E. Lubetzky, Y. Peres, Anatomy of a young giant component in the random graph, *Random Structures & Algorithms*, **39** (2011), 139-178.
- [13] R. Durrett. *Probability: Theory and Examples*, Vol. 3. Duxbury Press, 2004.
- [14] C. Greenhill, B. D. McKay, Random dense bipartite graphs and directed graphs with specified degrees, *Random Structures & Algorithms*, **35** (2009), 222-249.
- [15] C. Greenhill, B. D. McKay, Asymptotic enumeration of sparse nonnegative integer matrices with specified row and column sums, *Adv. in Appl. Math.*, **41** (2008), 459-481.
- [16] S. Janson, M. Łuczak. Asymptotic normality of the k -core in random graphs, *The annals of applied probability*, **18**, (2008), 1085-1137.
- [17] M. Karoński, T. Łuczak. The phase transition in a random hypergraph, *J. Comput. Appl. Math.*, **142**, (2002), 125-135.
- [18] R.M. Karp. The transitive closure of a random digraph, *Random Structures & Algorithms*, **1** (1990), 73-93.
- [19] R. M. Karp, M. Sipser. Maximum matchings in sparse random graphs, *Proceedings of the Twenty-Second Annual IEEE Symposium on Foundations of Computing*, (1981), 364-375.
- [20] D. Knuth, R. Motwani, B. Pittel. Stable husband, In *Proceedings of the first annual ACM-SIAM symposium on Discrete algorithms*, (1990), 397-404.
- [21] M. Krivelevich, B. Sudakov. The phase transition in random graphs: A simple proof, *Random Structures & Algorithms*, **43**, (2013), 131-138.
- [22] T. Łuczak. The phase transition in the evolution of random digraphs, *Journal of graph theory*, **14** (1990), 217-223.
- [23] T. Łuczak. Component behavior near the critical point of the random graph process, *Random Structures & Algorithms*, **1** (1990), 287-310.
- [24] T. Łuczak, T. Seierstad. The critical behavior of random digraphs, *Random Structures & Algorithms*, **35** (2009), 271-293.
- [25] B. McKay. Asymptotics for 0-1 matrices with prescribed line sums, *Enumeration and Design*, (Academic Press, 1984) (1984), 225-238.
- [26] B. McKay. Asymptotics for symmetric 0 – 1 matrices with prescribed row sums, *Ars Combin* **19** (1985), 15-25.
- [27] A. Nachmias, Y. Peres. Component sizes of the random graph outside the scaling window, *Alea* **3** (2007), 133-142.
- [28] X. Pérez-Giménez, N. Wormald. Asymptotic enumeration of strongly connected digraphs by vertices and edges, *Random Structures & Algorithms*, **43** (2013), 80-114.
- [29] B. Pittel. An urn model for cannibal behavior, *J Appl Probab*, **27** (1987), 522-526.
- [30] B. Pittel. On tree census and the giant component in sparse random graphs. *Random Structures & Algorithms*, **1** (1990), 311-342.

- [31] B. Pittel. Normal convergence problem? Two moments and a recurrence may be the clues, *Annals of Applied Probability*, **1**, (1999), 1260-1302.
- [32] B. Pittel. Counting strongly-connected, moderately sparse directed graphs, *Random Structures & Algorithms*, **43** (2013), 49-79.
- [33] B. Pittel, J. Spencer, N. Wormald. Sudden emergence of a giant k -core in a random graph, *J. Combinatorial Theory, Series B*, **67**, (1996), 111-151.
- [34] B. Pittel, R.S. Weishaar. The random bipartite nearest neighbor graphs, *Random Structures & Algorithms*, **15** (1999), 279-310.
- [35] B. Pittel, N. Wormald. Asymptotic enumeration of sparse graphs with a minimum degree constraint, *J. Combin. Theory Series A*, **101** (2003), 249-263.
- [36] B. Pittel, N.C. Wormald. Counting connected graphs inside-out, *Journal of Combinatorial Theory, Series B*, **93** (2005), 127-172.
- [37] T. Seierstad. On the normality of giant components, *Random Structures & Algorithms*, **43** (2013), 452-485.
- [38] V.E. Stepanov. Some features of connectedness of a random graph $G_m(t)$, *Theory Probab. Applies*, **15** (1970), 55-67.

A Appendix

A.1 Approximate Expectations

Here are the approximate expected values of the 15 pairwise products a^2, ab, \dots, k^2 . Note that each of these expected values is zero-degree homogeneous.

$$\begin{aligned}
\mathcal{E}_s[a^2] &:= \frac{\nu_i}{\nu_i + \nu_o} + \frac{\nu_o}{\nu_i + \nu_o} \frac{e^{z_i}}{e^{z_o} - 1} \frac{\nu_i z_i z_o}{(e^{z_o} - 1)\mu} \left(1 + \frac{\nu_i z_i z_o}{(e^{z_o} - 1)\mu}\right) \\
\mathcal{E}_s[b^2] &:= \frac{\nu_o}{\nu_i + \nu_o} + \frac{\nu_i}{\nu_i + \nu_o} \frac{e^{z_o}}{e^{z_i} - 1} \frac{\nu_o z_i z_o}{(e^{z_i} - 1)\mu} \left(1 + \frac{\nu_o z_i z_o}{(e^{z_i} - 1)\mu}\right) \\
\mathcal{E}_s[ab] &:= \frac{\nu_i}{\nu_i + \nu_o} \frac{e^{z_o}}{e^{z_o} - 1} \frac{\nu_o z_i z_o}{(e^{z_i} - 1)\mu} + \frac{\nu_o}{\nu_i + \nu_o} \frac{e^{z_i}}{e^{z_i} - 1} \frac{\nu_i z_i z_o}{(e^{z_o} - 1)\mu} \\
\mathcal{E}_s[ar_i] &:= \frac{\nu_i}{\nu_i + \nu_o} \frac{e^{z_o}}{e^{z_o} - 1} \frac{\nu - \nu_i - \nu_o}{e^{z_i} - 1} \frac{z_i z_o}{\mu} \\
\mathcal{E}_s[ar_o] &:= \frac{\nu_o}{\nu_i + \nu_o} \frac{e^{z_i}}{e^{z_i} - 1} \frac{\nu_i z_i z_o}{(e^{z_o} - 1)\mu} \frac{(\nu - \nu_i - \nu_o) z_i z_o}{(e^{z_o} - 1)\mu} \\
\mathcal{E}_s[ak] &:= \frac{\nu_i}{\nu_i + \nu_o} \frac{z_o e^{z_o}}{e^{z_o} - 1} + \frac{\nu_o}{\nu_i + \nu_o} \frac{e^{z_i}}{e^{z_i} - 1} \frac{\nu_i z_i z_o}{(e^{z_o} - 1)\mu} (1 + z_i) \\
\mathcal{E}_s[br_i] &:= \frac{\nu_i}{\nu_i + \nu_o} \frac{e^{z_o}}{e^{z_o} - 1} \frac{\nu_o z_i z_o}{(e^{z_i} - 1)\mu} \frac{(\nu - \nu_i - \nu_o) z_i z_o}{(e^{z_i} - 1)\mu} \\
\mathcal{E}_s[br_o] &:= \frac{\nu_o}{\nu_i + \nu_o} \frac{e^{z_i}}{e^{z_i} - 1} \frac{\nu - \nu_i - \nu_o}{e^{z_o} - 1} \frac{z_i z_o}{\mu} \\
\mathcal{E}_s[bk] &:= \frac{\nu_i}{\nu_i + \nu_o} \frac{e^{z_o}}{e^{z_o} - 1} \frac{\nu_o z_i z_o}{(e^{z_i} - 1)\mu} (1 + z_o) + \frac{\nu_o}{\nu_i + \nu_o} \frac{z_i e^{z_i}}{e^{z_i} - 1} \\
\mathcal{E}_s[r_i^2] &:= \frac{\nu_i}{\nu_i + \nu_o} \frac{e^{z_o}}{e^{z_o} - 1} \frac{\nu - \nu_i - \nu_o}{e^{z_i} - 1} \frac{z_i z_o}{\mu} \left(1 + \frac{\nu - \nu_i - \nu_o}{e^{z_i} - 1} \frac{z_i z_o}{\mu}\right) \\
\mathcal{E}_s[r_i r_o] &:= 0 \\
\mathcal{E}_s[r_i k] &:= \frac{\nu_i}{\nu_i + \nu_o} \frac{e^{z_o}}{e^{z_o} - 1} \frac{\nu - \nu_i - \nu_o}{e^{z_i} - 1} \frac{z_i z_o}{\mu} (1 + z_o) \\
\mathcal{E}_s[r_o^2] &:= \frac{\nu_o}{\nu_i + \nu_o} \frac{e^{z_i}}{e^{z_i} - 1} \frac{\nu - \nu_i - \nu_o}{e^{z_o} - 1} \frac{z_i z_o}{\mu} \left(1 + \frac{\nu - \nu_i - \nu_o}{e^{z_o} - 1} \frac{z_i z_o}{\mu}\right) \\
\mathcal{E}_s[r_o k] &:= \frac{\nu_o}{\nu_i + \nu_o} \frac{e^{z_i}}{e^{z_i} - 1} \frac{\nu - \nu_i - \nu_o}{e^{z_o} - 1} \frac{z_i z_o}{\mu} (1 + z_i) \\
\mathcal{E}_s[k^2] &:= \frac{\nu_i}{\nu_i + \nu_o} \frac{z_o e^{z_o}}{e^{z_o} - 1} (1 + z_o) + \frac{\nu_o}{\nu_i + \nu_o} \frac{z_i e^{z_i}}{e^{z_i} - 1} (1 + z_i)
\end{aligned}$$

A.2 Converting Gaussian Limits from $D(n, m)$ to $D(n, p)$

The proof of the following lemma is essentially a copy of Lemma 2 in Pittel [30], which allowed transfer of asymptotic normality results in $G(n, m)$ to $G(n, p)$.

Lemma A.1. *Let $k \geq 1$ be fixed, and let $\mathbf{Y} = \mathbf{Y}(D) \in \mathbb{R}^k$. Suppose that there exists a k -dimensional vector function $\boldsymbol{\mu} = \boldsymbol{\mu}(x) (x > 0)$ and a $k \times k$ symmetric matrix function $\mathbf{A} = \mathbf{A}(x)$ such that for $D(n, m = c_n n)$ and any $c_n \rightarrow c \in (0, \infty)$, $[\mathbf{Y} - n\boldsymbol{\mu}(c_n)]n^{-1/2}$ is asymptotically Gaussian with the zero vector of means and the covariance matrix $\mathbf{A} =$*

$\mathbf{A}(c)$. Suppose also that $\boldsymbol{\mu}(x)$ is continuously differentiable, and $\mathbf{A}(x)$ is continuous. Then, for $D(n, p = c_n/n)$, $(\mathbf{Y} - n\boldsymbol{\mu}(c_n))n^{-1/2}$ is also asymptotically Gaussian, with zero mean and covariance matrix

$$\mathcal{A} = \mathcal{A}(c) = \mathbf{A}(c) + c\boldsymbol{\mu}'(c)\boldsymbol{\mu}'(c)^T.$$

Proof. Let $\mathbf{u} \in \mathbb{R}^k$. We need to show that for $\boldsymbol{\mu}_n = \boldsymbol{\mu}(c_n)$, $\mathbf{A} = \mathbf{A}(c)$, $\mathcal{A} = \mathcal{A}(c)$,

$$\varphi_p(\mathbf{u}) := E_p[\exp(i\mathbf{u}^T(\mathbf{Y} - n\boldsymbol{\mu}_n)n^{-1/2})] \rightarrow \exp\left(-\frac{1}{2}\mathbf{u}^T\mathcal{A}\mathbf{u}\right), \quad (\text{A.1})$$

where $E_p[\circ]$ is the expectation over $D(n, p = c_n/n)$. Since $D(n, p)$ conditioned on the value of present edges $e(D(n, p)) = m$ is equal in distribution to $D(n, m)$, we have that

$$\varphi_p(\mathbf{u}) = \sum_{m=0}^{n(n-1)} \binom{n(n-1)}{m} p^m (1-p)^{n(n-1)-m} E_m[\exp(i\mathbf{u}^T(\mathbf{Y} - n\boldsymbol{\mu})^{-1/2})],$$

where $E_m[\circ]$ is the expectation over $D(n, m)$.

Fix $a > 0$, and we split the sum above into Σ_1 with those m such that $|m - n(n-1)p| \leq a\sqrt{n(n-1)p(1-p)}$, and Σ_2 with those remaining m . By the central limit theorem,

$$\begin{aligned} |\Sigma_2| &\leq \sum_{|m - n(n-1)p| > a\sqrt{n(n-1)p(1-p)}} \binom{n(n-1)}{m} p^m (1-p)^{n(n-1)-m} \\ &\rightarrow (2/\pi)^{1/2} \int_a^\infty e^{-z^2/2} dz \quad (n \rightarrow \infty). \end{aligned} \quad (\text{A.2})$$

Further, for all m in the sum Σ_1 ,

$$|m/n - c_n| = O(n^{-1/2}).$$

Note that

$$\begin{aligned} E_m[\exp(i\mathbf{u}^T(\mathbf{Y} - n\boldsymbol{\mu})^{-1/2})] &= \exp(i\mathbf{u}^T[\boldsymbol{\mu}(m/n) - \boldsymbol{\mu}]n^{1/2}) \\ &\quad \times E_m[\exp(i\mathbf{u}^T[\mathbf{Y} - n\boldsymbol{\mu}(m/n)]n^{-1/2})]. \end{aligned}$$

This second factor approaches $\exp(-\frac{1}{2}\mathbf{u}^T\mathbf{A}\mathbf{u})$, while the first term is

$$\exp\left(i(c_n)^{1/2}[\mathbf{u}^T\boldsymbol{\mu}'(c_n)]\frac{m - n(n-1)p}{\sqrt{n(n-1)p(1-p)}}\right) + o(1),$$

uniformly for m in Σ_1 . Therefore,

$$\begin{aligned} \Sigma_1 &\rightarrow (2\pi)^{-1/2} \int_{|z| \leq a} \exp(i(c)^{1/2}[\mathbf{u}^T\boldsymbol{\mu}'(c)]z - z^2/2) dz \\ &\quad \times \exp\left(-\frac{1}{2}\mathbf{u}^T\mathbf{A}\mathbf{u}\right). \end{aligned} \quad (\text{A.3})$$

Letting $a \rightarrow \infty$ in (A.2) and (A.3), we get that

$$\varphi_p(\mathbf{u}) \rightarrow \exp \left(-\frac{1}{2}c[\mathbf{u}^T \boldsymbol{\mu}'(c)]^2 - \frac{1}{2}\mathbf{u}^T \mathbf{A} \mathbf{u} \right),$$

which is equivalent to (A.1). □